NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ANALYZING THE EFFECTS OF HUMAN PERFORMANCE UNDER STRESS

by

Daniel B. Ammons-Moreno
Kathleen E. Pauls

June 2008

Thesis Advisor:	Samuel E. Buttrey
Second Reader:	David W. Meyer

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** June 2008	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Analyzing the Effects of Human Performance Under Stress			**5. FUNDING NUMBERS**
6. AUTHOR(S) Daniel B. Ammons-Moreno, Kathleen E. Pauls			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE**

13. ABSTRACT (maximum 200 words)

In order to analyze the effects of stress on human performance, we examined baseball players because of the large body of data and many measures of performance available. Clutch hitting is examined because a baseball player batting in a clutch situation is analogous to a person who is performing in a stressful situation. The more important, or clutch, the situation the more stress the player may feel. Statistical measures were used to determine if a player is able to perform better than his average ability in situations defined as clutch. Three different clutch definitions were used to examine eight consecutive years of baseball data. Major League Baseball (MLB) data showed an overall clutch effect; this was corrected for with a parameter, alpha, is specific to the definition of clutch. Once each player's non-clutch average minus the clutch average is corrected for with alpha, the chi-squared test is used to examine those differences. This analysis is also performed on the quartile values for batters who were ranked according to their difference, corrected by alpha. There is no evidence to support the claim that there are certain batters who perform better in clutch situations (compared to their own performance in non-clutch situations) than other batters.

14. SUBJECT TERMS Baseball, clutch hitting, binomial proportion, sign test			**15. NUMBER OF PAGES** 83
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

ANALYZING THE EFFECTS OF HUMAN PERFORMANCE UNDER STRESS

Daniel B. Ammons-Moreno
Ensign, United States Navy
B.S., United States Naval Academy, 2007

Kathleen E. Pauls
Ensign, United States Navy
B.S., United States Naval Academy, 2007

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED SCIENCE
(OPERATIONS RESEARCH)

from the

NAVAL POSTGRADUATE SCHOOL
June 2008

Author: Daniel B. Ammons-Moreno
Kathleen E. Pauls

Approved by: Samuel E. Buttrey
Thesis Advisor

David W. Meyer
Second Reader

James N. Eagle
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In order to analyze the effects of stress on human performance, we examined baseball players because of the large body of data and many measures of performance available. Clutch hitting is examined because a baseball player batting in a clutch situation is analogous to a person who is performing in a stressful situation. The more important, or clutch, the situation the more stress the player may feel. Statistical measures were used to determine if a player is able to perform better than his average ability in situations defined as clutch. Three different clutch definitions were used to examine eight consecutive years of baseball data. Major League Baseball (MLB) data showed an overall clutch effect; this was corrected for with a parameter, alpha, is specific to the definition of clutch. Once each player's non-clutch average minus the clutch average is corrected for with alpha, the chi-squared test is used to examine those differences. This analysis is also performed on the quartile values for batters who were ranked according to their difference, corrected by alpha. There is no evidence to support the claim that there are certain batters who perform better in clutch situations (compared to their own performance in non-clutch situations) than other batters.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

To analyze the affects of stress on human performance, this analysis focused on baseball players, because so much data is available. Clutch hitting is examined because of the measure's similarity to performance under stress. The more important, or clutch, the situation the more stress the player may feel. The extent to which a situation is "clutch" is described by factors such as runners in scoring positions, the number of outs, score differential, and the game inning. The situation can only be described as clutch if the batter is aware the situations importance to the overall game.

Statistical measures are used to study clutch hitting to determine if a player is able to perform better than his average ability in situations defined as clutch. Three different clutch definitions are used to examine eight consecutive years of baseball data. Each player has a known batting average in the non-clutch situations and a known batting average in the clutch situations. Using these two averages a difference is computed and examined under the three different definitions. A parameter, alpha, was calculated from the mean of the differences. Alphas were also generated for the different situations to see if there is a situational affect. Specific alphas were created for each situation, but simulation suggested that the model was not improved by specifying different alphas for different situations. An overall clutch effect was found. The more strict the clutch definition is, the larger the corresponding alpha. All of the alphas were found to be positive. This implies that on the whole the general population of batters tends to perform worse in clutch situations than their average performance.

Once each player's non-clutch average minus his clutch average is corrected for by alpha, the chi-squared test is used to examine those differences. There were two types of analysis done with the chi squared test. First, the data was used to create a binomial table. In this form there are five different combinations of negative ones and positive ones. The chi squared test was then performed on this binomial table. Second, the data was used to create a sign table which was tested with the chi squared test. This sign table

contains 16 different combinations of positives and negatives. Unlike the binomial table where "+−−+" is the same as "−−++", the sign table distinguishes the two outcomes.

A further examination of the chi-squared tests described earlier showed that the analysis was neglecting an interesting and surprisingly large bias. This bias was great enough to compromise any inferences that could be drawn from these tests. A method for determining an individual clutch effect that was unaffected by this bias was devised. The new method places batters into quartiles based on how much better each batter's clutch performance is then his non-clutch performance. The quartile placements are determined by how the batters compare to one another.

A league-wide clutch performance trend was observed. Several test verified that the distribution of clutch batting averages is different than the distribution of non-clutch batting averages when looking at all players. After establishing the general effect and correcting for it, no individual effect could be found. In sum, there is no evidence to support the claim that there are certain batters who perform better in clutch situations when compared to their performance in non-clutch situations than other batters.

ACKNOWLEDGMENTS

The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at 20 Sunset Rd., Newark, DE 19711.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

When a baseball player is called a clutch hitter, a reference is being made to a player's ability to hit better in certain situations. However, there is controversy over whether or not clutch ability exists. Do some batters hit better in certain situations because these situations are "clutch," or can these occasions where batters seem to perform abnormally well in certain situations be explained by probability? More generally, do certain people perform better or worse than their average performance in stressful situations? There are two main problems in trying to answer this question. First, how does one measure a person's average performance and measure the departure from that average performance in that stressful situation? Second, what defines a stressful situation? It is likely that there are situations that are stressful to most people, but presumably there could be situations that are stressful to certain individuals and not others. Furthermore, the idea that a situation is either stressful or not is an over-simplification of reality; a person can experience a range of stress. Baseball players are subject to differing amounts of stress throughout a season and their performances are constantly analyzed and documented.

Baseball players are ideal test subjects for the question at hand because their performance is quantifiably measurable and many years of baseball data is easily accessible. A person who performs better than his or her average performance in stressful situations is similar to a batter who makes an important play in a stressful batting situation. This type of batter is commonly referred to as a clutch hitter and therefore, examining the existence of clutch hitting is akin to answering the question of "do certain people perform better or worse than their average performance in stressful situations?"

B. LITERARY REFERENCE

In order for a player to be clutch, his performance needs to be in some way predictable. Grabiner (2006) formed specific situational definitions and then measured the performance in clutch and non-clutch situations. The difference in these two values is what Grabiner calls the clutch performance.[1] To measure the clutch performance, the expected wins were computed from both the raw data and from situational data. The probabilities of a win are computed before the batter steps up to the plate and again after the batter bats. The difference in the two measures is what Grabiner refers to as the clutch performance.

Others have also attempted to measure clutch hitting in a probabilistic fashion. Sauer and Hakes define a clutch situation to be one in where the impact of the player's performance on the probability of a victory is greater than that same performance in a normal situation.[2] The authors use a method which compares a player's productivity across different situations. The situation is said to be "key" if the probability impact of the play is twice as high as normal. Key situations encompass 10.9% of all the plate appearances. The situation is "meaningless" if the probability impact of the play is less than one quarter that of a normal play. These "meaningless" situations account for 16.0% of the plate appearances.

The probabilistic approach is flawed when attempting to answer questions about human performance under stress in that it requires that the outcome of the at-bat be known in order to determine whether the situation is clutch. When trying to determine whether a player's performance in certain situations is clutch it doesn't make sense to use an approach that requires the situation to depend on the performance of the player. Fuld's problem with the probabilistic approach is that although clutch is defined, there is still an arbitrary line drawn placing everything on one side clutch and everything on the other

---

[1] David Grabiner, Do Clutch Hitters Exist? (paper presented at the SABRBoston Presents Sabermetrics conference, May 20, 2006).

[2] Jahn H. Hakes and Raymond D. Sauer, "Are Players Paid for 'Clutch' Performance?" John E. Walker Dept. of Economics, Clemson University. Preliminary Draft (2003), http://people.albion.edu/jhakes/pdfs/clutch.pdf.

side not clutch.[3] The placement of that line can have huge impacts on the result. Fuld also expressed the need for separate measures of performance and importance. For example, a team is down by two in the ninth inning, there are two outs and there are two runners in scoring position. The upcoming batter needs to hit a home run to win. Hakes and Sauer call the situation important if the batter hits a home run and not very important if he strikes out. However, if the batter is just a bad batter he is likely to strike out every time. The fact that the batter is bad should not change the fact that the situation is important.

Cramer discussed the need for a measure of hitting timeliness and a measure of hitting quality. Cramer referenced the Harlan and Eldon Mills book, "Player Win Averages," which discussed how the brothers devised a measure that used the probable outcome of a baseball game. These probabilities were determined by computer play based on the average level of hitting for almost every one of the 8000 possible situations, such as two outs, runners on 1st and 2nd, tie game, top of the 6th, etc. Each game participant in every season is given "Win" or "Loss" points for how much his involvement increased or decreased the chances of the his team winning.[4] These points per player are accumulated to form the "Player Win Average" (PWA). There is also a Batter Win Average (BWA) that measures the quality of hitters. Cramer devised a formula that would compute the number of runs a league would have scored if a particular player were replaced by an average hitter. The difference in the two league run totals reflects the batter's average skills in producing runs for his team. This study compared players over a two-year period. The probabilistic problem is again seen in this study. Regardless of what the outcome of a batter's plate appearance is, the extent to which the situation is a clutch one should be unchanged.

Fuld approaches the problem of clutch hitting from another angle. Fuld used a regression on the hypothetical performances of a player against the importance of the situation. The "importance index" is independent of player performance and is used to

3 Elan Fuld, "Clutch and choke Hitters in Major League Baseball: Romantic Myth or Empirical Fact." 1st Draft (2005).

4 Richard D. Cramer, "Do Clutch Hitters Exist?" Baseball Research Journal (1977), http://www.geocities.com/cyrilmorong@sbcglobal.net/CramerClutch2.htm.

measure the inherent importance of the situation.[5] This index is calculated by determining how much the probability of winning the game would be altered by the current batters performance. The index measures only the importance of the situation to winning that particular game. The regression is done on the scatter plot with the importance index on the X axis and the on-base percentage plus slugging percentage (OPS) on the Y axis. The OPS has values from zero to five: 0-out, 1- walk/hit by pitch, 2-single, 3-double, 4-triple, 5- home run. This regression aims at finding the batters who hit better at important points in the game and identifies those that are good as "clutch" and those that are not as "choke".

The idea of using regression is appealing, but the creation of an arbitrary index raises some questions. It is hard to tell how accurate the importance index is. The index is based roughly on how helpful the batter's at-bat just was. Also the OPS scale goes from 0-5 with a single being better then a walk by one, a double being better than a single by one, etc. It is agreed upon that the home run is the best and that an out is the worst, but it is unclear by how much better each of these indices are from each other. A potential flaw in the index is that the value of each outcome increases linearly; an out counts as a zero whereas a single counts as a one. It may not always be the case that the value of a double exceeds that of a single by precisely the amount by which a single's value exceeds that of an out.

To analyze the affects of stress on human performance, eight consecutive years of data is analyzed to observe trends in players over all years. As in previous studies, measures are created to allow for the study of individual players rather than the study of the general population.

---

[5] Elan Fuld, "Clutch and Choke Hitters in Major League Baseball: Romantic Myth or Empirical Fact." 1st Draft (2005).

II. ANALYSIS

The first problem in answering the question, “Does clutch hitting exist?” is that there is no explicit definition of clutch that is universally accepted. In general, clutch hitting is when a batter performs uncharacteristically well in a stressful situation. There are many factors that would put stress on a batter, such as batting during a close game, batting with runners in scoring position, batting with one or two outs, the batter facing the minor leagues due to prior poor performance, and batting in an away game. Some of these factors are easier to search for than others. Sauer and Hakes (2003) state that the “clutch” of a situation is dependent upon how significant the outcome of the batter’s at bat in on the final outcome of the game. Since clutch hitting is merely a vehicle for the large question about performance under stress, the definition of a clutch situation used in this analysis must be limited to the factors that the batter currently sees. This is why Sauer and Hakes’ definition is not acceptable for our analysis. Additionally, there are many factors that would impact a batter’s stress level that are difficult to incorporate into the model. Such things include batters worrying about being demoted to the minor leagues, batters facing left- or right-handed pitchers, night or day games, and home or away games. These factors are left out of this analysis because of the difficulty in incorporating these in to the model. However, if it is the case that ignoring these other factors obscures our analysis so much that we cannot prove the existence of clutch hitting, then presumably the clutch effect is not very significant to the overall performance.

The definitions of clutch used in this paper include easily-measured game states: inning, score differential, runners on base, and number of outs. The batter is always aware of these game states so these are all reasonable factors that could stress a batter. The outcome of the plate appearance, positive or negative, does not influence the fact that the situation is clutch; a situation is classified as clutch based on the current status of the game before the batter faces his first pitch. This classification scheme is slightly naïve

because game state can change between pitches as in the case of a stolen base. This analysis will ignore the situations that transformed from non-clutch to clutch during an individual plate appearance because this does not happen very many times during a year.

There are three different definitions of a clutch situation used in this paper. The first definition (Def1) of a clutch situation classifies clutch situations as the set of all plate appearances that occur in the seventh inning or later with runners in scoring position and a score differential less than or equal to three. A runner in scoring position is when there is a runner on or past second base. This definition was chosen first because in general, people feel that during the last few innings of a game is when the situations become more clutch. However, not all plate appearances late in a game are clutch. For example, if a team is winning by a substantial amount then there is less pressure on the batters of either team to make a big play than there would be if it were a close game. The number of clutch situations that occurred in the year 2003 according to this definition was 10,573. The average number of clutch situations for all eight years is approximately 10,746.

The second definition (Def2) provides a loose definition of clutch. For this definition, a situation is clutch if the game is in the fifth inning or later, there are one or more runners in scoring position and the score differential is less than or equal to four. This is a looser definition and therefore more batters experienced clutch situations than in the first definition. The number of clutch situation seen in 2003 was 21,457. The average number of clutch situations for all eight years is approximately 21,802.

The third definition (Def3) is the most restrictive and would be viewed as clutch by any reasonable standard. This definition requires the game be in the seventh inning or later, with runners in scoring position, a score differential less than or equal to three, and two outs. The number of the clutch situations seen the 2003 with this definition are 4,946. The code used to change these definitions in SPLUS is located in Appendix A. The average number of clutch situations for all eight years is approximately 4,955. Table 1 highlights all the attributes of each definition.

	Inning	Runners/ Scoring Pos.	Outs	Score Diff.	Avg/Yr
Def1	≥ 7	Yes	Any	≤ 3	10,746
Def2	≥ 5	Yes	Any	≤ 4	21,802
Def3	≥ 7	Yes	2	≤ 3	4,955

Table 1. Definition table.

Batters that are found to perform above average (that is, whose clutch averages exceed their non-clutch ones) according to these definitions in a given year would likely be called clutch hitters. This analysis will search for the specific batters who perform above average in these clutch situations year after year. However, probabilistically, out of all the batters in the major leagues there should be some that perform above average year after year just due to random chance. Therefore, the proof of clutch hitting, and ultimately the proof of deviations in average performance for people under stress, would be determined by the presence of a statistically significant number of batters who perform above average in clutch situations over many years.

A. DATA

1. The Need for an Alpha

The data used in the analysis of clutch comes from the last eight consecutive years, the 2000-2007 seasons. The data was provided by Retrosheet.[6]

	Id	Vis	Inn	TeamAtBat	O	B	S	OrigSeq	VSc	HSc	Batter	PHand	BHand	Event	Type	BatEvt	AB	Hit	SH	SF	OPlay	RBI
1	ANA200303300	TEX	1	0	0	3	2	BBBCCFB	0	0	gland001	R	R	W	14	T	F	0	F	F	0	0
2	ANA200303300	TEX	1	0	0	1	1	BFX	0	0	everc001	L	R	S5/BG5S.1-2	20	T	T	1	F	F	0	0
3	ANA200303300	TEX	1	0	0	0	1	CX	0	0	rodra001	R	R	5(2)3/GDP.1-2	2	T	T	0	F	F	2	0
4	ANA200303300	TEX	1	0	2	3	2	BSBBSX	0	0	gonzj002	R	R	S/78.2-H	20	T	T	1	F	F	0	1
5	ANA200303300	TEX	1	0	2	3	1	BCBBB	1	0	palmr001	L	R	W.1-2	14	T	F	0	F	F	0	0
6	ANA200303300	TEX	1	0	2	1	2	CSBS	1	0	sierr001	L	R	K	3	T	T	0	F	F	1	0
7	ANA200303300	TEX	1	1	0	1	2	CFBX	1	0	ecksd001	R	R	S8	20	T	T	1	F	F	0	0
8	ANA200303300	TEX	1	1	0	1	0	BX	1	0	erstd001	L	R	26(1)/FO	2	T	T	0	F	F	1	0
9	ANA200303300	TEX	1	1	1	0	0	X	1	0	salmt001	R	R	S8.1-2	20	T	T	1	F	F	0	0
10	ANA200303300	TEX	1	1	1	0	2	CCX	1	0	andeg001	L	R	9/F9.2-3	2	T	T	0	F	F	1	0

Figure 1. Sample of the events from the 2003 plate appearances.

[6] The information used here was obtained free of charge from and is copyrighted by Retrosheet. Interested parties may contact Retrosheet at 20 Sunset Rd., Newark, DE 19711.

Figure 1 lists ten of the 187,449 plate appearance that occurred in the year 2003. The first statistic to be analyzed is the difference between the non-clutch hits divided by the number of non-clutch plate appearances minus the number of clutch hits divided the number of clutch plate appearances for each player:

$$Difference = \frac{nonclutch\ hits}{nonclutch\ plate\ appearances} - \frac{clutch\ hits}{clutch\ plate\ appearances} \qquad [1]$$

Using this statistic (clutch difference statistic), a data frame is generated to analyze the distribution of these differences for each player in 2003. The clutch definition used to generate this data frame comes from Def1.

Batter ID	clutch.hits	clutch situations	non.clutch hits	non.clutch situations	Difference
abada001	0	2	2	17	0.1176
aberb001	0	3	2	34	0.0588
abreb001	11	41	162	654	-0.0206
alfoe001	6	34	127	552	0.0536
allec001	0	1	5	24	0.2083
almoe001	0	2	26	109	0.2385
alomr001	6	36	127	562	0.0593
aloms001	4	11	48	193	-0.1149
aloum001	8	33	150	605	0.0055
ameza001	1	6	21	114	0.0175

Figure 2. The first ten batters who have at least one clutch plate appearance in the year 2003 under Def1.

Figure 3 shows the distribution of differences for each batter calculated by using Equation 1.

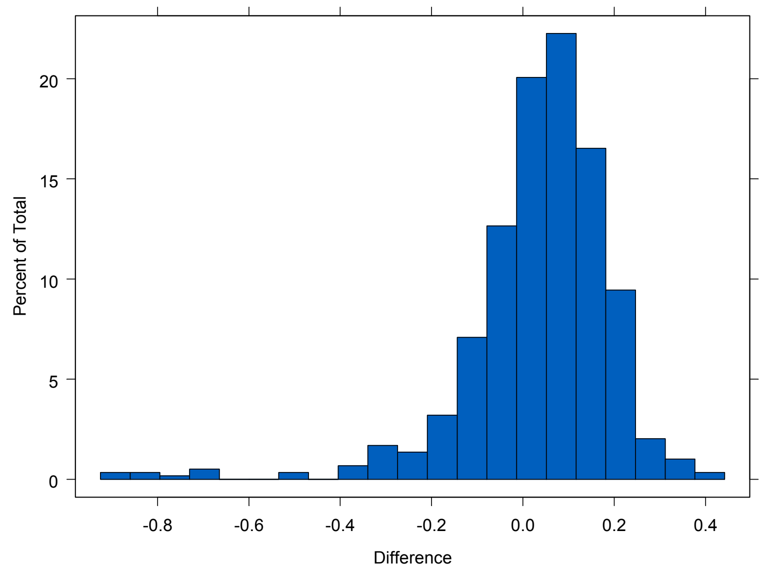

Figure 3. Histogram of the clutch difference statistic for players with at least one clutch plate appearance in the year 2003.

As seen in Figure 3, the distribution seems to be centered to the right of zero. This indicates that on the average, more Major League Baseball players perform worse in clutch situations than they do in non-clutch situations; this phenomenon is known as "choking" and can be seen as the opposite of clutch hitting. The values near negative one are caused by players who have a very small number of plate appearances; these players' differences distort the overall shape of the histogram. Figure 4 shows the same set of differences, restricted to batters with 20 or more clutch plate appearances.

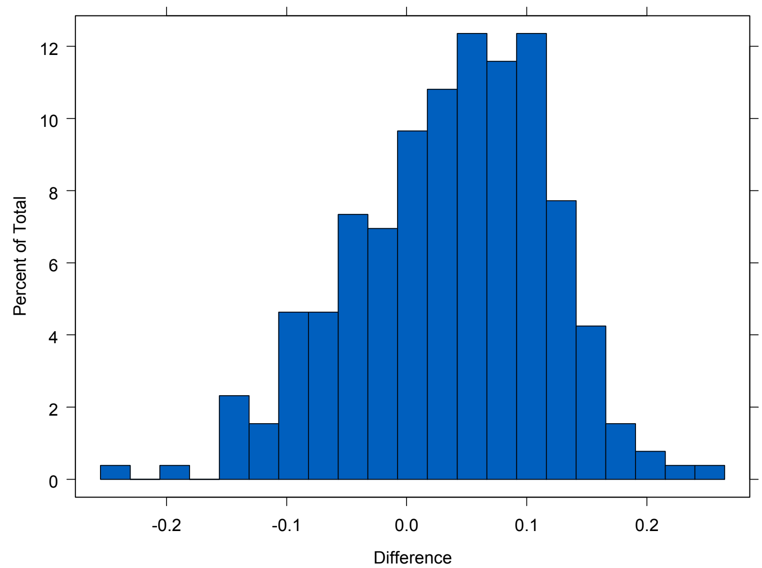

Figure 4. Histogram of the clutch difference statistic for players with at least 20 clutch plate appearances in the year 2003.

The histogram is centered off to the right of zero. Therefore, the difference between the non-clutch hits divided by the number of clutch situations minus the number of clutch hits divided by the number of clutch situations is primarily positive.

A simple two-sided t-test will not accurately test the hypothesis that the mean of the difference distribution is zero. Each player in the data frame has a different number of clutch and non-clutch plate appearances; players with large numbers of plate appearances should have a larger impact on the t statistic than other players. Standardizing each player's difference by dividing each clutch difference statistic by the standard deviation of that difference would create a new statistic that would be properly weighted by each player's number of plate appearances. Assuming the probability each player gets a hit in either a clutch or non-clutch situation is a Bernoulli trial with probability of success equal to that player's "true" clutch or non-clutch batting average, then the variance of the difference is equal to the sum of the variance of the two binomial distributions. The standardized clutch difference statistic is calculated using this formula:

$$
\begin{aligned}
&Var\left(\frac{c_1}{n_1}-\frac{c_2}{n_2}\right)\\
&=Var\left(\frac{c_1}{n_1}\right)+Var\left(\frac{c_2}{n_2}\right)\\
&=\frac{\left(\frac{c_1}{n_1}\right)\left(1-\frac{c_1}{n_1}\right)}{n_1}+\frac{\left(\frac{c_2}{n_2}\right)\left(1-\frac{c_2}{n_2}\right)}{n_2} \qquad [2]
\end{aligned}
$$

Variables c_1 and c_2 are the number of non-clutch hits and clutch hits. Variables n_1 and n_2 are the number of non-clutch situations and clutch situations. The difference is then divided by square root of the variance. Figure 5 shows the histogram for the standardized differences.

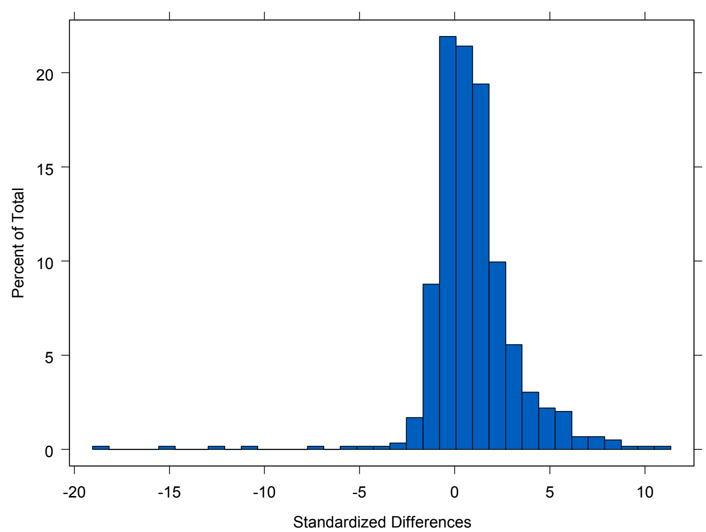

Figure 5. Histogram of the standardized differences for plate appearances in 2003 with batters who had one or more clutch situations

The process of standardizing the differences should result in approximately equal variances; assuming that the standardized differences seen in Figure 5 are normally distributed and that the differences are independent of one another, then a two-sided one sample t-test can be performed on the standardized differences. The two-sided one sample t-test results in a t statistic of 9.6102 on 592 degrees of freedom; this yields a p-

value of zero. Given the p-value is zero the null hypothesis is unlikely to be true and there must be some difference between the two ratios that comprise the clutch difference statistic.

This result could be an aspect of the fact that the standardized statistic includes walks, sacrifice bunts, hit-by-pitch, and sacrifice flies. It is possible that these plays could happen in significantly different proportions for clutch and non-clutch situations because of strategy on the part of either the batting team's manage or the pitcher. This could create an imbalance among the non-clutch and clutch averages that would skew our findings. For example, team managers might order batters to bunt more often when the game is close and there is a man on third. This would dramatically affect the clutch difference statistic computed because sacrifice bunts do not count as hits and batters are being told to bunt more often in clutch situations. Clutch is the batter's ability to perform well in stressful situations and being told to bunt by a team manager should not count against a batter. Walks sometimes happen as a strategic decision made by a pitcher and it could be true that walks occur in different proportions for clutch and non clutch situations; for the same reason as before, walks should not impact the measurement of a batter's clutch ability. In the standardized statistic, decisions by the pitcher and the team manager are not removed, so they do impact the current batter's clutch difference statistic. Since the goal is to measure the batter's clutch ability, strategic decisions made by external actors should not impact the batter's clutch difference statistic.

One subset of plate appearances is at-bats. In baseball, an at-bat is any plate appearance that does not result in a walk, hit-by-pitch, sacrifice hit, or sacrifice fly. This will be the subset that will be used for the analysis' continuation. The standardized clutch difference statistic that is now being examined is the same as before except that situations that resulted in walks, bunts, hit-by-pitch, and sacrifice flies have been removed entirely. This new difference is exactly the difference between non-clutch batting averages and clutch batting averages; since batting averages are computed only from at-bats. Using clutch Def1, a new histogram (Figure 8) is generated to show the distribution of these differences for the year 2003 after restricting the analysis to at-bats.

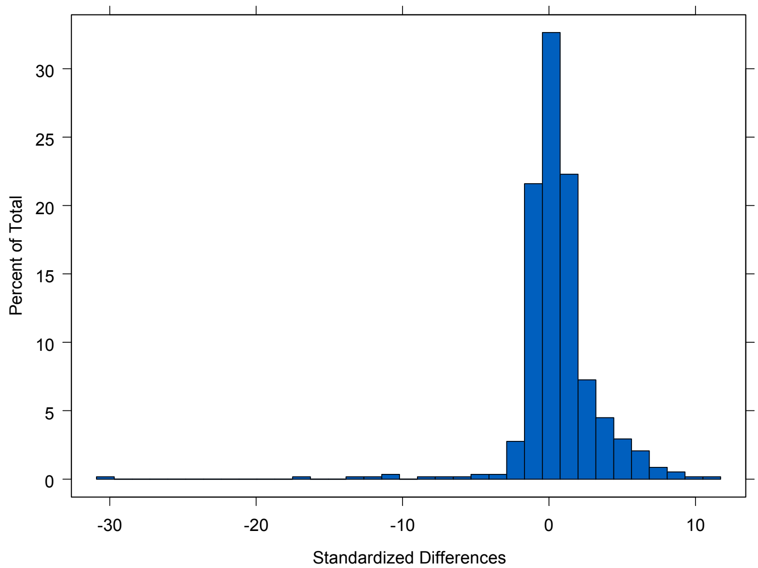

Figure 6. Histogram of the standardized clutch difference statistic for players who had at least one clutch at bat in the year 2003

Once again, in order to test the hypothesis that the distribution of the clutch difference statistic has a mean of zero, the differences needs to be standardized. This accounts for the varying number of at-bats each batter has in the year 2003. Using the same standardization formula in Equation 2 on the differences, a new t-test can be executed to test the hypothesis. The two-sided one sample t-test results in a t statistic of 6.522 on 538 degrees of freedom; this yields a p-value of zero. This low p-value implies the true mean of this standardized difference is not zero. This result only applies to the differences that originated in the year 2003; ultimately, eight recent consecutive years of major league baseball data is available and combining all the years allows for a more powerful result.

Figure 7 is a portion of the table of batters who have at least one clutch at-bat in any of the eight consecutive years of available data.

Batter ID	clutch.hits	clutch situations	non.clutch hits	non.clutch situations	Difference	Variance	Standard Difference
abada001	0	1	2	16	0.1250	0.0068	1.5119
abboj002	7	17	63	240	-0.1493	0.0151	-1.2165
abbok002	0	7	34	150	0.2267	0.0012	6.6306
aberb001	22	59	190	809	-0.1380	0.0042	-2.1334
aberr001	1	16	68	315	0.1534	0.0042	2.3667
abreb001	66	232	1310	4396	0.0135	0.0009	0.4444
abret001	4	11	41	155	-0.0991	0.0223	-0.6639
acevj002	0	1	2	42	0.0476	0.0011	1.4491
adamr002	16	46	198	818	-0.1058	0.0052	-1.4731
agbab001	10	43	208	757	0.0422	0.0044	0.6354

Figure 7. A portion of the table of batters who have at least one clutch at-bat in any of the eight years of consecutive data.

Figure 8 contains the histogram created from the standardized differences shown in Figure 7.

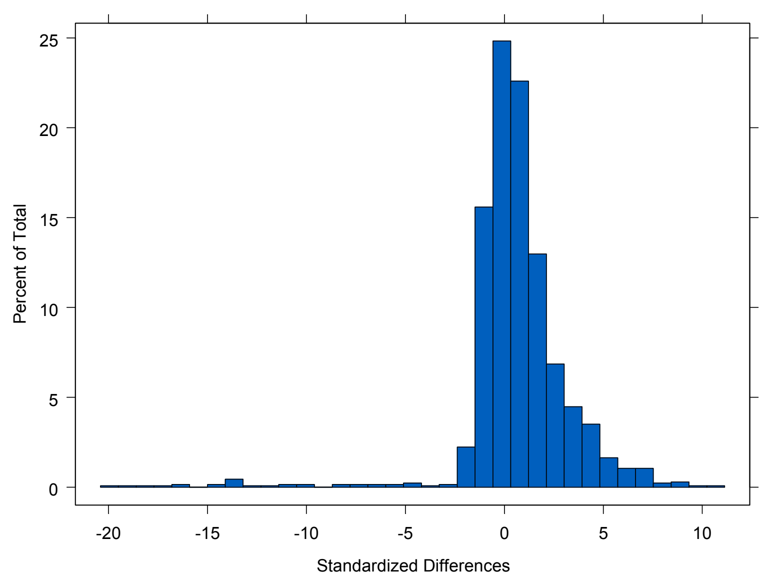

Figure 8. Histogram for the standardized differences of players with one or more clutch at bats in a given year, summed over eight years.

A t-test is then run on the standardized differences for the players who had one or more at bat in at least one of eight years. The two-sided one sample t-test results in a t statistic of 8.8288 on 1340 degrees of freedom; this yields a p-value of zero. The null hypothesis is that the mean of this standardized distribution is zero; given the low p-value

and the fact that this test is conducted on eight years of data, it is likely that on the average, major league batters perform differently in non-clutch situations than they do in clutch situations. Looking at the t-tests for Def2 and Def3, the p-values are shown to be zero.

clutch definition	t-statistic	p-value	degrees of freedom
Def1	8.829	0	1340
Def2	7.746	0	1572
Def3	12.456	0	1238

Table 2. Table of t-test results for all definitions of clutch for batters with one or more at bats summed over the years 2000-2007.

As the clutch definition becomes more restrictive the t statistic becomes larger. This could imply that the more difficult the clutch situation the worse the batter's performance. The results for all three definitions further suggest that on the whole the batters perform differently in non-clutch situations than they do in clutch situations. However, while the general trend is interesting a more interesting discovery would be to find evidence that certain batters have inherent clutch ability. The question is to find out if there are people who can perform better or worse in clutch situations, not whether the general population performs better or worse.

2. Analyzing Alpha

Given that there is a difference on the whole between batting averages between non-clutch and clutch situations, it makes sense to correct for the general effect in order to examine the individual player performance differences. The correcting factor that describes the overall difference between these non-clutch and clutch at-bats will be known as alpha. Alpha is be calculated by summing all the non-clutch hits and dividing that by the sum of all the non-clutch at-bats then subtracting the sum of all the clutch hits divided by the sum of the clutch situations. These sums come from the eight year table comprised of unique batters who had at least one clutch at-bat in at least one of the eight years. The alpha calculated based on each definition is shown in Table 3.

Definition	Alpha
Def1	0.0128
Def2	0.0011
Def3	0.0344

Table 3. The alphas calculated for each definition.

The alphas shown in Table 3 are the mean differences between the mean non-clutch batting average and the mean clutch batting average for this subset of players over the years 2000-2007. Once again the trend that was visible among the t statistic is also visible among these alphas; as the clutch definition ranges from least severe (Def2) to most severe (Def3) the value of the alpha corresponding to the definition increases. Since alpha is always positive the clutch batting average is always lower than the non-clutch batting average and as alpha increases the difference between the two averages becomes even larger. The non-clutch hits and non-clutch at-bats corresponding to batters who never had a clutch at-bat are ignored in the computation of these alphas.

The latter approach for determining a single alpha given a clutch definition could be naive. Ruane states that, "…batters do no hit equally well in all situations."[7] Ruane exhibits batting averages that differ depending on the number of outs and the position of any runners. Furthermore, it is generally accepted that it is easier for a batter to get on base in certain situations; for example, if there is a runner on first then typically the first baseman must play closer to first base. The tight first baseman position leaves more of the infield open, giving the batter a greater area in which to hit safely. By Ruane's definition, a "situation" is the current state of the game when the batter steps up to the plate. For example, one out with runners on first and third is an example of a situation. There are 24 combinations of outs and runner positions. If batting averages are fundamentally different for different situations then perhaps the clutch effect might be different, as well, requiring different alphas for differing situations.

There are two possibilities being considered; either there is one alpha that describes the grand clutch effect across all situations or there is a different clutch effect

[7] Tom Ruane, "In Search of Clutch Hitting," Baseball Research Journal (2005), http://retrosheet.org/Research/RuaneT/clutch_art.htm..

for each situation. The latter possibility calls for multiple alphas. Since the alphas may differ, an alpha is calculated for each situation by subtracting the mean clutch batting average for that situation from the mean non-clutch batting average for that situation. This operation yields an alpha for every situation for which a comparison is possible for each definition of clutch. For example, Def1 and Def2 do not have a specification for the number of outs in order for a situation to be considered clutch, but Def3 requires two outs for a situation to be considered clutch. This means that Def3 allows for only 6 alphas where Def2 and Def1 allow for 18 each. At-bats are grouped into the 24 situations and flagged as either clutch or non-clutch. Then, the outcome of each at-bat is recorded and the non-clutch and clutch batting averages are computed for each situation. In Figure 9, the situational batting average table is shown for Def1.

Situation	ABNon	ABClu	HitNon	HitClu	OutNon	OutClu	AvgNon	AvgClu	Alpha
1.0	71711	NA	20909	NA	50802	NA	0.2916	NA	NA
1.1	85888	NA	24574	NA	61314	NA	0.2861	NA	NA
1.2	84771	NA	22500	NA	62271	NA	0.2654	NA	NA
12.0	14822	3123	4150	837	10672	2286	0.2800	0.2680	0.0120
12.1	27098	7629	7280	1908	19818	5721	0.2687	0.2501	0.0186
12.2	34370	9359	8338	2156	26032	7203	0.2426	0.2304	0.0122
13.0	4920	1163	1751	398	3169	765	0.3559	0.3422	0.0137
13.1	10342	2637	3569	875	6773	1762	0.3451	0.3318	0.0133
13.2	15705	4081	4029	1029	11676	3052	0.2565	0.2521	0.0044
2.0	18762	4059	5082	1028	13680	3031	0.2709	0.2533	0.0176
2.1	30796	8367	8042	1986	22754	6381	0.2611	0.2374	0.0238
2.2	38672	9499	9647	2236	29025	7263	0.2495	0.2354	0.0141
23.0	3332	661	1071	200	2261	461	0.3214	0.3026	0.0189
23.1	7012	1553	2206	456	4806	1097	0.3146	0.2936	0.0210
23.2	9390	2324	2239	517	7151	1807	0.2384	0.2225	0.0160
3.0	2548	606	818	221	1730	385	0.3210	0.3647	-0.0437
3.1	9234	2159	3173	717	6061	1442	0.3436	0.3321	0.0115
3.2	16020	3800	3871	877	12149	2923	0.2416	0.2308	0.0108
Empty.0	333766	NA	89546	NA	244220	NA	0.2683	NA	NA
Empty.1	236236	NA	60449	NA	175787	NA	0.2559	NA	NA
Empty.2	185389	NA	46957	NA	138432	NA	0.2533	NA	NA
Loaded.0	3813	992	1258	340	2555	652	0.3299	0.3427	-0.0128
Loaded.1	8686	2915	2775	911	5911	2004	0.3195	0.3125	0.0070
Loaded.2	12468	3977	3097	994	9371	2983	0.2484	0.2499	-0.0015

Figure 9. Situational batting average table for clutch definition one.

The leftmost column names the situation. The numbers before the decimal indicate the runner position (Empty meaning that there are no runners and Loaded

meaning runners on first, second and third) and the number after the decimal gives the number of outs. ABNon and ABClu contain the number of non-clutch at-bats and clutch at-bats. The HitNon and HitClu columns contain the number of hits in non-clutch and clutch situations. AvgNon and AvgClu are the calculated batting averages for each of the situations. The Alpha column is just the difference between the AvgNon and the AvgClu averages; these alphas are the situational alphas estimated from the data. Notice that in Figure 9 only 18 of the 24 alphas have numerical values. This is because six of the 24 situations never produce clutch at-bats under Def1. Now the question is, "Are these alphas significantly different from each other, or could one grand alpha have created the individual alphas?" In other words, could there be an overall alpha that applies to all situations and the reason the individual alphas appear to be different from each other is random chance? Or, could it be that each situation has a different alpha, implying that each situation has a different effect on clutch at-bats? In order to answer this question a satisfactory grand alpha must first be computed. In Table 3, three different alphas for the different definitions are shown. These alphas are one possible set of grand alphas that correspond to each definition of a clutch situation. Table 4shows alphas computed from Figure 9. These alphas, unlike those in Table 3, include players with no clutch at-bats.

Definition	Grand Alpha
Def1	0.0098
Def2	0.0004
Def3	0.0304

Table 4. The grand alphas calculated for each definition.

The alphas measure the difference the between the non-clutch and the clutch batting averages. For players who never have a clutch at-bat, how their performance would be different in a clutch situation is unknown. The assumption is that this given player's performance would change by this factor, grand alpha. If Table 4 alphas were used, a player whose clutch abilities that were not measured would be allowed toinfluence alpha. For this reason, the alphas in Table 3 will be used to see if the situational alphas are necessary or if the one grand alpha for each definition from Table 3 is sufficient.

After having chosen the grand alphas, SPLUS can be used to simulate the clutch and non-clutch batting averages for each situation using the different situational non-clutch batting averages and the grand alpha we selected for each definition. For example, using the table in Figure 9, the SPLUS simulation would use the non-clutch batting average, AvgNon, for the runners on first and second with zero outs, situation 12.0, to simulate 14,822 at-bats, ABNon, and then simulate 3,123 clutch at bats, ABClu, using the situational non-clutch batting average, AvgNon, minus the grand alpha corresponding to Def1 from Table 2. For convenience, the row for situation 12.0 shown in Figure 10 and a portion of simulated alphas for this situation are shown in Figure 11. The alpha shown in Figure 10 is the actual situational alpha estimated from the data, not a simulated one.

situation	ABNon	ABClu	HitNon	HitClu	OutNon	OutClu	AvgNon	AvgClu	Alpha
12.0	14822	3123	4150	837	10672	2286	0.2800	0.2680	0.0120

Figure 10. Section of Figure 9 used in example of the simulation.

0.0205	0.0107	0.0579	0.0175	0.0120	0.0038	0.0128	0.0040	0.0108	0.0138

Figure 11. Ten simulated alphas for the 12.0 situation (runners on first and second with no outs) under Def1.

The assumption this simulation is attempting to test is whether or not the observed situational alphas could have arisen from just the non-clutch situational batting average and one general correcting factor, grand alpha. If the simulated alphas cover the same range as the real alphas then the simulation has shown that one alpha can be used to create the different alphas seen in Figure 9. The simulation was run 10,000 times and the standard deviation of the simulated alphas was greater than the standard deviation of the estimated alphas 524 times. This shows that roughly 5% of the time the simulated alphas are more varied than the real alphas. When applied to the other definitions, the simulation again yielded standard deviations that were greater than those of the real alphas approximately 5% of the time. While the simulation is not a perfect representation of the real alphas, it appears to be close enough to argue in favor of the claim that a single alpha could have created the alphas shown in Figure 9. Therefore, the grand alphas in Table 3 will be used as the correcting factor when searching for specific batters who fare better or worse in clutch situations.

THIS PAGE INTENTIONALLY LEFT BLANK

III. RESULTS

A. CHI SQUARED ANALYSIS

The general effect for the Major Leagues in clutch situations can now be corrected for with the adequate correction factor called alpha. Now the analysis can search for individuals who perform better in clutch situations than in non-clutch situations. The analysis will now examine a new statistic, corrected difference, shown in Equation 3.

$$\text{Corrected Difference} = \text{Nonclutch Batting Average} - \text{Clutch Batting Average} - \text{alpha} \qquad [3]$$

The corrected difference associated with each batter can be computed for each year in which the batter had at least one clutch at-bat.

SPLUS can be used to apply the corrective factor to each players non-standardized clutch difference. Figure 12 shows a portion of the table of batters who had at least one clutch at-bat in the year 2003.

Batter ID	clutch.hits	clutch situations	non.clutch hits	non.clutch situations	Difference	Variance	Standard Difference	AlphaDiff	sign
abada001	0	1	2	16	0.1250	0.0068	1.5119	0.1122	1
aberb001	0	2	2	32	0.0625	0.0018	1.4606	0.0497	1
abreb001	11	30	162	547	-0.0705	0.0081	-0.7823	-0.0833	-1
alfoe001	6	25	127	489	0.0197	0.0077	0.2248	0.0069	1
allec001	0	1	5	23	0.2174	0.0074	2.5276	0.2046	1
almoe001	0	2	26	98	0.2653	0.0020	5.9489	0.2525	1
alomr001	6	25	127	491	0.0187	0.0077	0.2128	0.0059	1
aloms001	4	11	48	183	-0.1013	0.0221	-0.6818	-0.1141	-1
aloum001	8	27	150	538	-0.0175	0.0081	-0.1943	-0.0303	-1
ameza001	1	3	21	102	-0.1275	0.0757	-0.4633	-0.1402	-1

Figure 12. Portion of the table for batters with at least one clutch at-bat in the year 2003 under Def1.

The alpha that was applied to the Difference column to create the AlphaDiff column came from the alpha for Def1 in Table 2. AlphaDiff is the difference corrected by alpha. The sign column simply represents the sign of the AlphaDiff column. However, its significance is that a player with a negative sign is a player who performed better, in

2003, in clutch situations than in non-clutch situations after alpha had been taken into account. A sign column could be computed for each batter for each year. The signs from each year can be combined to make a larger sign matrix for all batters for all eight years. The sign matrix has eight columns, each corresponding to a year between 2000 and 2007; the matrix is filled only with negative ones, positive ones, and zeros. Zeros occur for batters who did not meet the number of required clutch situations for that given year. Figure 13 contains the first ten batters who had at least one clutch at-bat in at least one of the eight years.

Batter ID	year2000	year2001	year2002	year2003	year2004	year2005	year2006	year2007
abada001	0	0	0	1	0	0	0	0
abboj002	-1	-1	0	0	0	0	0	0
abbok002	1	0	0	0	0	0	0	0
aberb001	0	-1	-1	1	0	1	0	0
aberr001	0	0	0	0	0	0	1	1
abreb001	1	1	1	-1	1	0	-1	1
abret001	0	0	0	0	0	0	0	-1
acevj002	0	0	0	0	1	0	0	0
adamr002	0	0	0	0	-1	-1	-1	-1
agbab001	1	1	-1	0	0	0	0	0

Figure 13. Sign matrix of batters who had at least one clutch at-bat in at least one of eight years.

If it is the case that no batter has any inherent clutch ability, but that there is simply a general effect (alpha) for all batters in clutch situations, then the probability a player performs better in clutch situations than his non-clutch batting average minus alpha is fifty percent (but see section III.B). The non-clutch average minus alpha is the same as the clutch average, under the hypothesis that no player has inherent clutch ability. However, each player has an observed clutch average which we compute from the data. If it is the case that the player's observed clutch average is greater than the theoretical clutch average, then the player would have a negative one in the sign matrix for that year. For example, if there was a batter who outperformed his hypothetical clutch batting average for all eight years, i.e. had all negative ones in the sign matrix, then it would be safe to say that he, as an individual, has innate clutch ability. Still, there is a

chance that a batter like this could exist under the original assumptions. If clutch hitting was a real phenomenon, there would be an unusual number of batters with large numbers of negative signs.

One way to determine if there is individual clutch ability is to use a chi squared test. SPLUS was used to determine the number of times, over the course of eight years, a player performed better in clutch situations than his theoretical batting average. This number is a simple conditional sum and can be added to the matrix shown in Figure 13; now the number of times a player outperformed his theoretical clutch batting average can be easily seen in Figure 14.

Batter ID	year2000	year2001	year2002	year2003	year2004	year2005	year2006	year2007	Sums
abada001	0	0	0	1	0	0	0	0	0
abboj002	-1	-1	0	0	0	0	0	0	2
abbok002	1	0	0	0	0	0	0	0	0
aberb001	0	-1	-1	1	0	1	0	0	2
aberr001	0	0	0	0	0	0	1	1	0
abreb001	1	1	1	-1	1	0	-1	1	3
abret001	0	0	0	0	0	0	0	-1	1
acevj002	0	0	0	0	1	0	0	0	0
adamr002	0	0	0	0	-1	-1	-1	-1	4
agbab001	1	1	-1	0	0	0	0	0	2

Figure 14. Sign matrix of batters who had at least one clutch at-bat in at least one of eight years with a sums column.

Under the hypothesis that the probability a player performs better than his theoretical clutch batting average is fifty percent in each year, the expected distribution of these sums is known. The expected number of batters in each category *n*, i.e. 0, 1,…8, is equal to the total number of batters divided by the binomial probability of *n* successes in eight trials with a probability of success of 0.5. For example, the expected number of batters out of 300 who should outperform their theoretical clutch batting averages eight years in a row is 300 divided by 2^8, or 1.172 batters. However, very few batters appear in all eight years. Under the usual rules for application of the chi-squared test, all expected values are greater than or equal to five (Devore 2008, 507), only four years of data can be used.

The next problem occurs in dealing with batters who have only a small number of at-bats. For example, a player with one clutch at-bat will have a clutch batting average of one or zero for that given year. The null hypothesis is that there is a fifty percent chance that this better will perform better than his theoretical clutch batting average. In the case of this batter with one clutch at-bat, the probability that he performs better than his clutch batting average is not fifty percent. Given this batter's overall batting average is 0.25 then there is roughly a twenty-five percent chance that he will perform better than his clutch batting average and a seventy-five percent chance that he will not perform better. The larger problem here is an issue of granularity that causes bias. The issue of bias will be discussed in full detail in the following section. There are not enough clutch at-bats for these batters to get reasonable clutch batting averages. To avoid this problem the required number of at-bats for the clutch performance is set at 20 clutch at-bats per year. Figure 15 shows part of the larger sign table for batters who had at least 20 clutch at-bats in at least one of the eight years of data under Def1.

Batter ID	year2000	year2001	year2002	year2003	year2004	year2005	year2006	year2007
aberb001	0	-1	-1	0	0	0	0	0
abreb001	1	1	1	-1	1	0	-1	1
adamr002	0	0	0	0	0	-1	0	0
agbab001	1	0	0	0	0	0	0	0
alfoe001	1	1	0	1	1	-1	0	0
alfoe002	0	0	0	0	0	0	-1	0
alicl001	1	1	0	0	0	0	0	0
alomr001	1	-1	-1	1	0	0	0	0
aloum001	1	-1	1	-1	1	1	0	1
ameza001	0	0	0	0	0	0	0	1

Figure 15. Sign matrix of batters who had at least 20 clutch at-bats in at least one of eight years.

As seen in Figure 15, some batters see at least 20 clutch situations every year while others have only seen 20 clutch at-bats in one year. The new 20 at-bat restriction further reduces the number of batters who can meet the requirement and thus will increase the number of zeros in the table. This is another reason for reducing the size of the categories from eight to four years. After imposing the 20 clutch at-bat requirement, there are 189 batters who met the requirement in at least four years. Table 5 shows the number of batters who meet the 20 clutch at-bat requirement in 4,5,6,7, and 8 years.

Years	4	5	6	7	8
Counts	53	47	37	32	20

Table 5. Counts of batters for each category who met the 20 at-bat requirement in at least one of four year under Def1.

Here the category refers to the number of years in which the batters meet the at-bat requirement. Using the four year chi-squared analysis the expected number in each category is greater than five. For the batters who had more than four years of data, only the most recent four years were used. An additional problem is posed by players who have four years of at least 20 clutch at-bats but for whom those years are not consecutive. Eliminating these players would be extremely restrictive because most players have a few years where they did not achieve 20 clutch at-bats. The way this analysis will deal with this issue is to ignore the breaks and simply analyze the most recent four years with actual results for each batter. Table 6 is the chi-squared table for the 189 batters who met the clutch at-bat requirement.

Category	0	1	2	3	4
Observed	8	47	75	46	13
Expected	11.8125	47.25	70.875	47.25	11.8125

Table 6. The observed and expected table for batters who had more that 20 clutch at-bats and four years of data under Def1.

The category refers to the number of years a batter performed worse than his theoretical clutch batting average. The observed values match up closely to the expected values. The chi-squared goodness of fit test, calculated in SPLUS, results in a chi-squared statistic of 1.6243 and a *p*-value of 0.8044. Given the high *p*-value, there is no reason to disbelieve the null hypothesis that for any year, there is a fifty percent chance that a batter's true clutch batting average will be better than his non-clutch batting average corrected by alpha. In other words, apart from the league-wide clutch effect, intrinsic clutch ability does not appear to vary from batter to batter in a statistically significant

way. There are enough batters with at least 25 clutch at-bats to perform another chi-squared test. The second test results in a chi-squared statistic of 5.187 and a *p*-value of 0.269. There is still little evidence to support rejecting the null hypothesis for this clutch definition.

Under Def2, there are significantly more batters with 20 or more at-bats. This fact allows for higher clutch at-bat requirements, which will ultimately make for better resolution on the clutch batting averages. The observed table for batters with more than 20 clutch at-bats under Def2 is shown in Table 7.

Category	0	1	2	3	4
Observed	17	80	146	68	24
Expected	20.94	83.75	125.63	83.75	20.94

Table 7. The observed and expected table for batters who had more that 20 clutch at-bats and four years of data under Def2.

The chi-squared test performed on this table results in a *p*-value of 0.106. Different chi-squared tests can be performed with higher at-bats and Table 8 sums up the results of these tests.

# at-bats	p-value
20	0.1064
25	0.2680
30	0.0331
35	0.0986
40	0.1059
45	0.0623
50	0.0083

Table 8. Table of *p*-values for the different number of binomial at-bats under Def2.

The *p*-values in Table 8 fluctuate quite a bit as the clutch at-bat requirement is increased; however, two of these *p*-values are below 0.05, and on the whole, all of these *p*-values are fairly low. Although a couple *p*-values are below the 0.05 significance level, most are not. Furthermore, thus far in the analysis many significance tests have been performed. If the null hypothesis were true in all of these tests, it would still be expected

to see one or two of these tests result in low *p*-values. Therefore, the conclusion is that there is not enough evidence to reject the null hypothesis which supports an individual clutch ability.

Def3 eliminates too many batters for the 20 clutch at-bat requirement. The chi-squared test can be performed with the requirement eased to 12 clutch at-bats, but the resolution of the clutch batting average for batters with 12 clutch at-bats is poor. The results of the chi-squared test for batters with at least 12 clutch at-bats do not favor rejecting the null hypothesis. The *p*-value is 0.7103 and even if it were significantly lower, the poor resolution on the clutch batting averages would cast doubt on any significant conclusions drawn from such a test.

Another type of chi-squared test can be performed on this sign table. Instead of counting the number of years in which a player had a positive difference, a table can be made that breaks up the four most recent years of player differences into 16 different outcomes. For example, the earlier test places all players who had a positive difference of three in the same category, but in the new test, a player who has a positive difference for three years in a row followed by a negative difference his final year would be placed in a different category than a player who had two positive years followed by a negative and then followed by a positive. This makes a total of 16 different outcomes which means a minimum of 80 players with at least four years of 20 clutch at-bats or more is required for this test. The null hypothesis for this new test is that all the outcomes are equally likely. This is a chi-squared test to determine if the distribution of the 16 outcomes is uniform.

For the first clutch definition, the table of outcomes used in the chi-squared test is shown in Figure 16.

++++	+++-	++-+	++--	+-++	+-+-	+--+	+---	-+++	-++-	-+-+	-+--	--++	--+-	---+	----
13	11	10	18	10	13	6	13	15	11	12	10	15	9	15	8

Figure 16. Sign table of outcomes for batters in four years with at least 20 clutch at-bats in all four years according to Def1.

A chi-squared test performed on this table under the null hypothesis that the outcomes are all equally likely results in a chi-squared statistic of 11.89 on 15 degrees of

freedom for a p-value of 0.687. In this case, the hypothesis that all outcomes are equally likely is reasonable and there is no reason to reject it. There is a sufficient number of batters in this table such that the 20 at-bat requirement can be increased to 25. This allows for more resolution in the clutch batting averages and this also focuses the search for clutch ability more on players who have more clutch at-bats. This test results in a p-value of 0.225 and once again there is no reason to reject the null hypothesis that the outcomes are all equally likely.

For clutch definition 2, the table of outcomes for batters with more than 20 clutch at-bats is shown in Figure 17.

++++	+++-	++-+	++--	+-++	+-+-	+--+	+---	-+++	-++-	-+-+	-+--	--++	--+-	---+	----
24	19	12	27	19	23	19	23	18	24	26	22	27	18	17	17

Figure 17. Sign table of outcomes for batters in four years with at least 20 clutch at-bats in all four years according to Def2.

There are 335 batters in this table so it will be possible to increase the at-bat requirement. The chi-squared test on this table results in chi-squared statistic of 12.749 on 15 degrees of freedom for a p-value of 0.622. Again, there is no reason to reject the null hypothesis. The table of outcomes generated after increasing the at-bat requirement to 35 is shown in Figure 18.

++++	+++-	++-+	++--	+-++	+-+-	+--+	+---	-+++	-++-	-+-+	-+--	--++	--+-	---+	----
11	14	10	23	11	15	18	14	16	5	21	13	21	14	9	9

Figure 18. Sign table of outcomes for batters in four years with at least 35 clutch at-bats in all four years according to Def2.

There are 111 fewer batters in this table than there were in the Def2 20 clutch at-bat table, but this is still well over the eighty batter requirement. The chi-squared test performed for this table results in a chi-squared statistic of 26.143 for a p-value of 0.037. This low p-value casts doubt on the null hypothesis of equally likely outcomes and that would imply that some outcomes are favored. There are still plenty of batters, so the at-bat requirement can be increased further. The p-values for the different chi-squared tests are shown in Table 9.

# at-bats	p-values
20	0.6217
25	0.6720
30	0.0924
35	0.0366
40	0.0307
45	0.0488
50	0.0137

Table 9. Table of *p*-values for the different number of sign at-bats under Def2.

All of the *p*-values for tests performed with 35 or more clutch at-bats in Table 9 are less than 0.05. This implies that some of these outcomes are more likely than others. By looking at Figure 17 one can see which outcomes are favored. However, there appears to be no obvious reason why "++−−" is more likely than "−++−". Ultimately all outcomes not being equally likely implies that individual batters do not all have a fifty percent chance to outperform their theoretical clutch batting average for a given year.

For the strict definition of clutch, Def3, there were not enough batters that met the 20 clutch at-bat requirement in order to satisfy the "expected value of each outcome greater than or equal to five" rule of thumb for the chi-squared test. Easing the restriction from 20 clutch at-bats to 12 clutch at-bats results in 106 batters; this number is now enough to meet the chi-squared rule of thumb. The chi-squared test done on the new outcome table results in a *p*-value of 0.201. There is no reason to reject the null hypothesis at this point. Even if the *p*-value had been less than 0.05 and consequentially the null hypothesis was rejected, this would not be very informative; with only 12 clutch at-bats, many batters will have unrealistic clutch batting averages in comparison to their realistic non clutch batting averages which we taken from much larger numbers of non clutch at-bats.

The expected value rule of thumb can be bypassed using another rule provided by Conover. Conover states that for samples sizes greater than 10, and for analyses that involve three or more categories, a chi-squared test is acceptable as long as all the expected values are greater than .25 and as long as the sample size squared divided by the

number of categories is greater than or equal to ten.[8] Using the twenty at-bat restriction and Def3, there are too few batters even to use Conover's rule. There is only one batter who meets the 20 at-bat requirement in three years. Ultimately, Def3 is too restrictive for this analysis.

B. NOTE ON BIAS

The null hypothesis that the probability of any given batter outperforming his theoretical clutch batting average in any year is fifty percent is not in fact, exactly true. Let d be the "true" non-clutch batting average of a given batter. This analysis assumes that d is known because the non-clutch batting averages are estimated based on many observations (usually 200-400 non-clutch at-bats). Also assumed to be known is the theoretical clutch batting average because it is simply the non-clutch batting average minus alpha ($c = d - alpha$). Now let c' be the observed clutch batting average. Under the null hypothesis, the expected value of c' is c and c' is an unbiased estimator of c. However, this analysis uses the sign of ($c' - c$), and the hypothesis is that it is equally likely that this statistic will be positive or negative. The new question is, "is sign($c' - c$) an unbiased estimator of 0?"

Suppose a batter is observed with X clutch hits over n clutch at-bats. Then c' must take on one of the values $0/n$, $1/n, \ldots, n/n$; if c were exactly equal to one of these values, then sign($c' - c$) would be equal to 0. In this analysis, c is determined in part by alpha; since alpha is measured to a high degree of precision, it would be impossible for a batter with even 200 clutch at-bats to obtain an observed clutch batting average that was equal to his non-clutch batting average minus alpha. Therefore, also assume that c is not exactly equal to any of the values $0/n$, $1/n, \ldots, n/n$.

Let S = sign($c' - c$); then S is equal to one if $c' > c$ (this happens with probability equal to $\Pr(X/n > c) = \Pr(X > nc)$) and S equals negative one if $c' < c$ which occurs with $\Pr(X < nc)$. If X/n should happen to be exactly equal to c, the contribution to E[S] would of course be 0. Therefore, the expected value of S is shown in Equation 4.

[8] W.J. Conover, *Practical Nonparametric Statistics* (New York: John Wiley & Sons Inc, 1999), 241.

$$E[S] = (1)\ \Pr(X > nc) + (-1)\ \Pr(X < nc)$$
$$= (1 - \Pr(X < nc)) - \Pr(X < nc) = 1 - 2\ \Pr(X < nc) \qquad [4]$$

For a typical batter with a "true" clutch batting average of 0.268 and 20 clutch at-bats with five clutch hits, the bias that would be incurred in attempting to measure the sign($c' - c$) for that batter would be -0.0877 (0.268 is the overall batting average for all MLB batters in the year 2007). This bias is substantial and must be corrected for if the chi-squared tests performed in the previous section are to have any merit. Unfortunately, the number of clutch at-bats, the "true" clutch batting average, and the number of clutch hits each batter made each year determines how much each player's sign($c' - c$) is biased (the number of clutch hits will always be assumed to be equal to the player "true" clutch batting average multiplied by his number of clutch at-bats rounded down[9]). For example, given a player whose "true" clutch batting average is 0.268, the amount by which the bias affects the given player changes based on how many clutch at-bats the player in that year; this effect is shown in Figure 19.

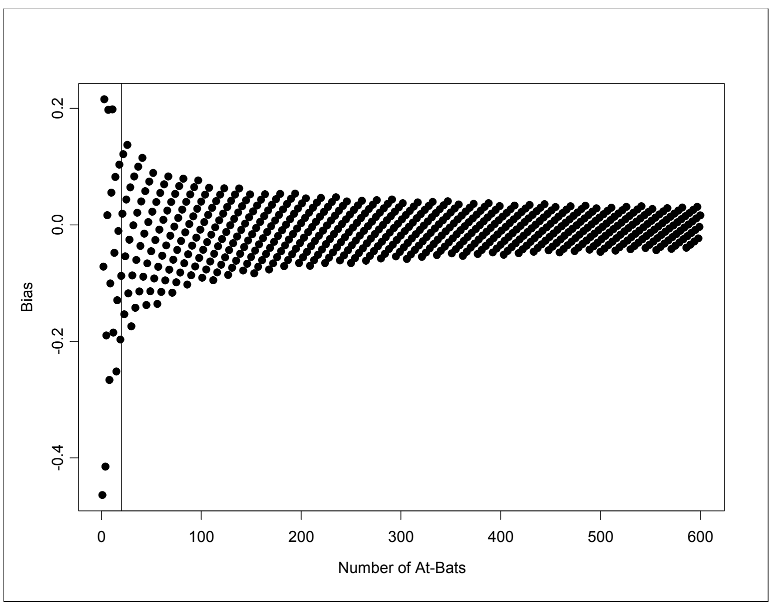

Figure 19. Plot of the bias for the varying number of at-bat requirements. The vertical line marks our at-bat requirement used shows the range of the bias at that requirement.

[9] The number of clutch hits a batter makes impacts the bias associated with measuring the sign($c' - c$) for the given batter. In order to simplify the exploration of the bias a very likely number of clutch hits a batter would make is that batter's batting average multiplied by the number of clutch at-bats.

The vertical line in Figure 19 is placed at "at-bats = 20." This shows how the much the bias could affect the determination of sign($c' - c$) for each player with a "true" clutch batting average of 0.268 and anywhere from 20 to 100 clutch at-bats (note that the bias is present even when the number of clutch at-bats is near 600.) However, not all players have "true" clutch batting averages equal to the league wide average. Figure 20 shows the how the bias changes based on a player's batting average provided that player had exactly 20 at-bats.

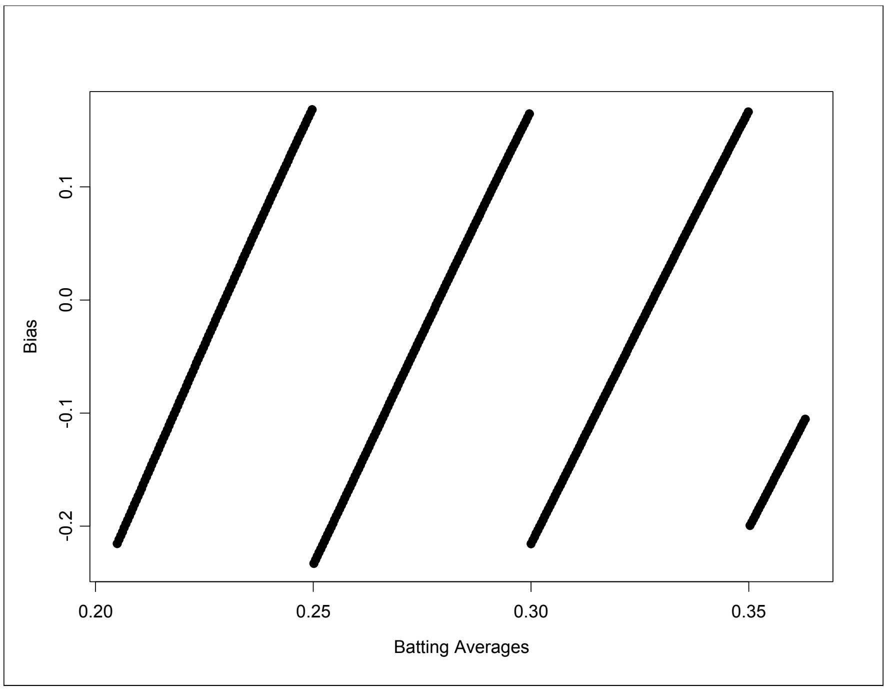

Figure 20. Bias shown for 1000 uniformly distributed "true" batting averages between 0.149 and 0.851 for batters who had at least 20 clutch at-bats.

The "jumps" from one line to the next line correspond to the precise values of batting averages that are possible to obtain with 20 clutch at-bats, i.e. 0.2, 0.25, 0.3… Figure 20 shows that the magnitude of the bias can be quite large across all batting averages for players with exactly 20 clutch at-bats. Figure 19 showed how the bias can be quite large for a specific batting average across a large number of clutch-bats. Finally, Figure 21 shows five box and whisker plots that each represent 1000 individuals with 20, 25, 30, 35, and 40 clutch-bats and a range of uniformly distributed clutch batting averages between 0.205 and 0.363.

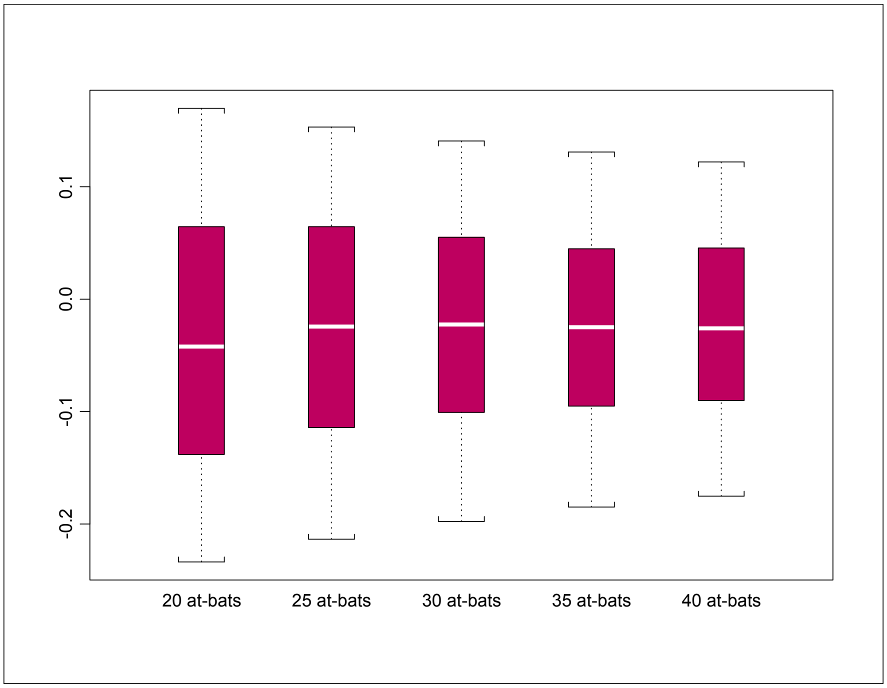

Figure 21. Each Box plot represents 1000 individuals with the corresponding number of at-bats and a range of batting averages uniformly distributed between 0.205-0.363.

The bias associated with measuring the sign($c' - c$) is both interesting and complex. Correcting each batter's sign($c' - c$) for each year is beyond the scope of this analysis and this will be mentioned in the further study section. However, there is a way to set-up a similar chi-squared test that does not rely on such heavily biased measurements as the previous tests.

C. CHI-SQUARED STANDARD QUARTILE SUMS

To avoid the impact of the bias created by the sign analysis, a new approach is taken. Rather than assign a sign value to the difference between the non-clutch batting average and the clutch batting average, the difference will be used directly. As before, the difference values are corrected by alpha and then standardized. The bias was created when the sign of the standardized values corrected by alpha were used; since the magnitude of each difference is now being considered, the previous bias is gone. In order for the batters to appear in the table they need to have at least 20 clutch at-bats in at least

four of the eight years. Since the four most recent years are used it is often the case that the standard difference with alpha for each batter comes from different years. For example, “batter A” might have eight years in which he fulfilled the clutch at-bat requirement, but “batter B” might only have fulfilled the clutch at-bat requirement in the first and last two of the eight years. This means that “batter B’s” standardized difference from the year 2000 will be compared to “batter A’s” standardized difference in the year 2003. Within each of the four years the batters are placed into quartiles depending upon how well each batter did when compared to how other batters performed that year. This is done for every batter in each year. If individual clutch hitting ability does not exist, then the probability that any individual batter will place in any of the quartiles, 1, 2, 3, or 4, is equally likely and independent from year to year. For example, if a given batter places in the first quartile in a given year, the probability that the batter places in the first quartile in the next year would still be 0.25 under that hypothesis. If it were the case that the given batter was more likely to place in the first quartile the following year, that would argue in favor of an individual clutch hitting ability. This example shows why the assumption of no individual clutch ability is analogous to the equally likely quartile placements from year to year. Figure 22 shows the first ten batters and the quartiles they were placed into for the four years most recent years that they had at least 20 at-bats in under Def1.

Batter ID	year 1	year 2	year 3	year 4	sums
abreb001	4	2	1	3	10
alfoe001	3	3	3	2	11
alomr001	4	1	1	3	9
aloum001	2	3	3	4	12
andeg001	2	1	1	4	8
aurir001	2	4	2	3	11
ausmb001	3	2	3	3	11
bagwj001	4	3	2	2	11
barrm003	2	1	2	2	7
batit001	3	2	1	2	8

Figure 22. Portion of the batters who appeared in the standardized difference with alpha table, then ranked and summed under Def1.

The expected values are calculated by the probability (under the null hypothesis) that an individual obtains a given sum over four years of quartile rankings multiplied by

the total number of individuals that appeared in the table. The probability that an individual obtains a given sum can be calculated by analyzing the total number of ways a batter can achieve each possible sum value. The numbers of possible combinations for each sum value are shown in Table 10.

Quartile Sums	4	5	6	7	8	9	10	11	12	13	14	15	16
Permutations	1	4	10	20	31	40	44	40	31	20	10	4	1

Table 10. Table of the different permutations for producing the quartile sum values.

For example, there is one way for a batter to achieve a sum of four; the batter had to have been in the first quartile all four years to achieve a sum of four. Similarly there is only one way to achieve a sum of 16. There are four ways to achieve a quartile sum of five. The batter could have been in the first quartile three of the four years and then in the second quartile the last year. There are four permutations of 1,1,1,2. Once the total number of permutations for each sum are known, the probability that an individual batter will achieve a given sum is the number of permutations for that sum divided by 256 (256 is the total number of permutations across all sums, $4\times4\times4\times4$); this is because under the assumption that individual clutch ability does not exist, each permutation is equally likely.

If each player is equally likely to appear in any of the quartiles for a given year the expected quartile for that year is the probability of any quartile multiplied by the quartile value. This yields an expected quartile value of 2.5 for any given year. Over all four years the expected sum for any given batter meeting the clutch at-bat requirement is ten. The expected numbers of batters for each quartile sum and the actual number of batters for each quartile sum are shown for each definition.

Quartile Sums	4	5	6	7	8	9	10	11	12	13	14	15	16
Observed	1	1	5	17	26	32	30	27	23	14	9	3	1
Expected	0.74	2.95	7.38	14.77	22.89	29.53	32.48	29.53	22.89	14.77	7.38	2.95	0.74

Table 11. Table of observed and expected values for the number of batters in the quartile sums under Def1.

Quartile Sums	4	5	6	7	8	9	10	11	12	13	14	15	16
Observed	3	4	12	21	46	54	64	47	35	23	16	9	1
Expected	1.31	5.23	13.09	26.17	40.57	52.34	57.58	52.34	40.57	26.17	13.09	5.23	1.31

Table 12. Table of observed and expected values for the number of batters in the quartile sums under Def2.

Quartile Sums	4	5	6	7	8	9	10	11	12	13	14	15	16
Observed	0	2	2	18	20	29	34	26	22	11	9	1	0
Expected	0.68	2.72	6.80	13.59	21.07	27.19	29.91	27.19	21.07	13.59	6.80	2.72	0.68

Table 13. Table of observed and expected values for the number of batters in the quartile sums under Def3.

Under the null hypothesis, the distribution of the quartile sums should be distributed as shown in Tables 11, 12, and 13 for each definition. Figure 23 is the histogram for the sums found for batters under Def1 who had at least 20 clutch at-bats in four consecutive years. Figure 24 is the histogram for the sums found for batters under Def2 who had at least 20 clutch at-bats in four consecutive years. In order to calculate the sums for Def3 the clutch at-bat requirement was lowered to ten. The histogram for batters' quartile sums under Def3 is shown in Figure 25.

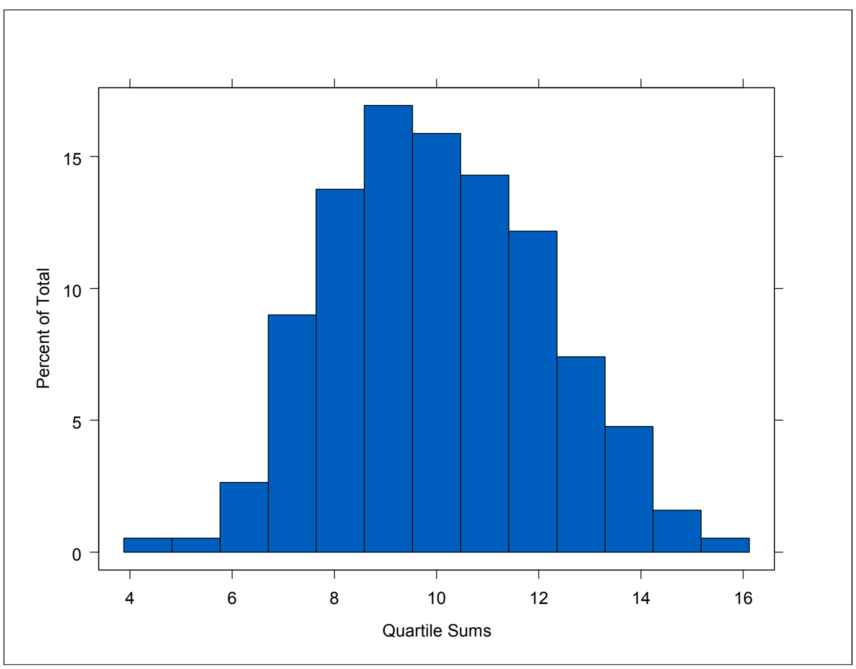

Figure 23. Histogram for the quartile sums for Def1.

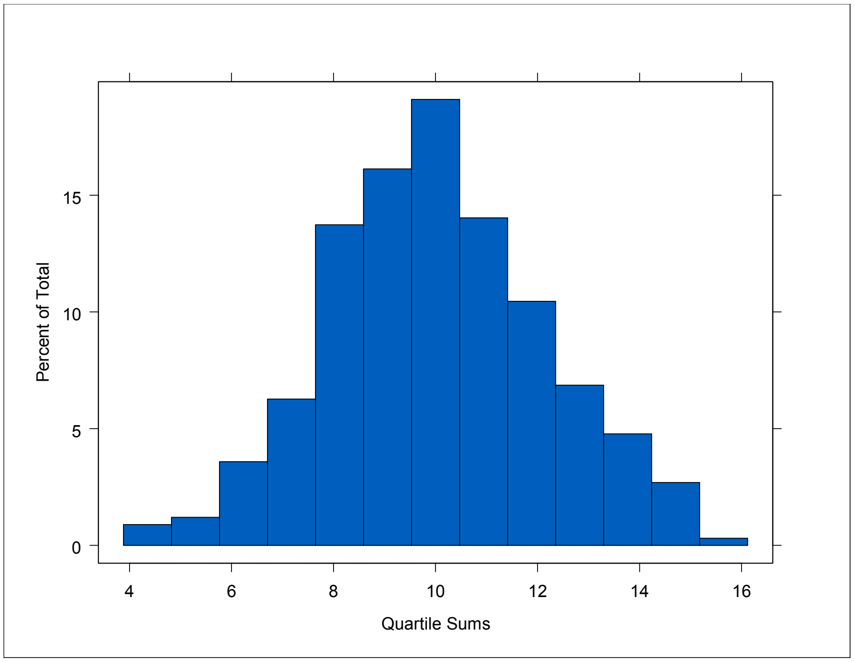

Figure 24. Histogram for the quartile sums for Def2.

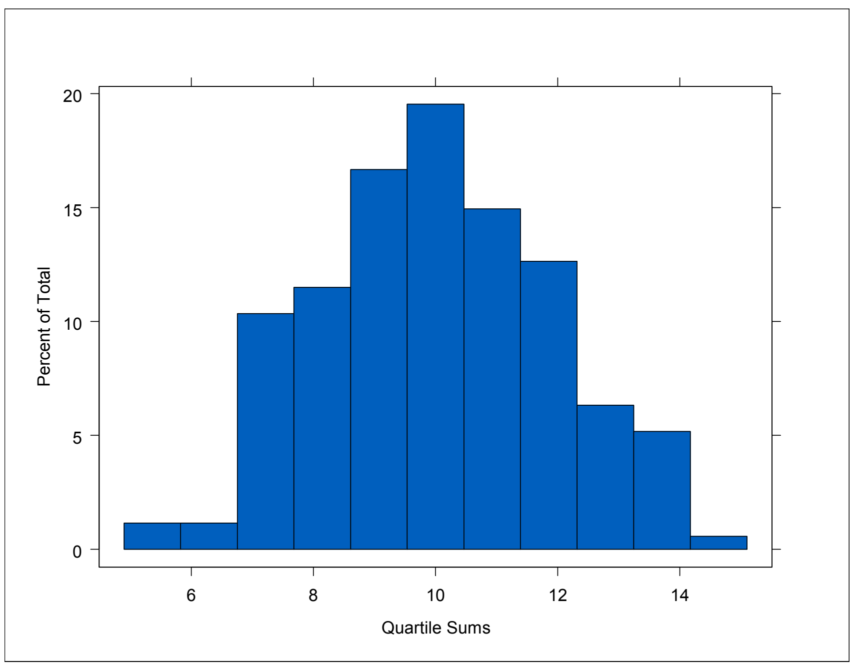

Figure 25. Histogram for the quartile sums for Def3.

These histograms are approximately symmetrical and centered on ten. This is what is expected under the null hypothesis. A chi-squared test can be performed to

determine if the observations made could have arisen from the expected distributions that have been calculated under the assumption that individual clutch hitting ability does not exist.

The chi-squared test will be performed under Conover's rules (covered earlier) seeing as how the expected values shown in Tables 10, 11 and 12 fall below 5 in places. However, all three tests meet Conover's criteria for chi-squared tests. The results of the three chi-squared tests, one for each clutch definition, performed on the summed quartiles are shown in Table 14.

Definition	p-values
Def1	0.9831
Def2	0.5975
Def3	0.6609

Table 14. Table of *p*-values for chi-squared analysis done on the quartile sums for each definition.

As seen in Table 14, the *p*-values are all significantly higher than .05. These results give no reason to reject the null hypothesis that there is no individual clutch ability.

D. SIMULATION

Another way to test the null hypothesis is by simulation. Under the null hypothesis, the probability that any given player achieves a specific quartile sum is shown in Table 15.

Quartile Sums	4	5	6	7	8	9	10	11	12	13	14	15	16
Probabilty	0.004	0.016	0.039	0.078	0.121	0.156	0.172	0.156	0.121	0.078	0.039	0.016	0.004

Table 15. Table of probabilities for a player achieving a specific quartile sum.

These probabilities are simply the total number of outcomes shown in Figure 15 divided by 256. S-Plus can be used to generate a sum for each player using the probabilities; if the sums generated resemble the actual sums measured, then there would be no reason to reject the null hypothesis. On the other hand, if clutch hitting were real the observed distribution would be more spread out than the hypothetical since players

with persistent clutch ability would be in the first quartile unusually often. Out of 10,000 simulations, the number of times the generated sums had a greater standard deviation than the actual sums was 4137 for Def1. The simulations for Def2 and Def3 also yielded similar results showing that there is not enough evidence to refute the assumptions used to generate these sums. Therefore, there is no evidence that the null hypothesis is wrong.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. CONCLUSIONS AND FURTHER STUDY

A. CONCLUSIONS

This analysis's goal was to examine human performance under stress. The popular idea of clutch hitting in baseball correlates very well with the idea of human performance under stress. The first examination of the data showed the existence of a Major League-wide difference between clutch hitting and non-clutch hitting. This trend was observable under each definition of clutch and ultimately several t-tests proved that the distribution of clutch batting averages was not the same as the distribution of non-clutch batting averages. In fact, for each definition, the corrective factor alpha was always measured to be positive. This implies that clutch performance is worse than non-clutch performance in general. The value of alpha increases as one shifts from loose to strict definitions of clutch and although this analysis did not address the fact that some situations are more stressful than others, it is not unreasonable to suggest that the clutch situations in the strict definition are more stressful on average than the clutch situations in the loose definition. Because the alphas become larger as the clutch definition becomes more strict, this could imply that clutch batting performance becomes worse as the situations become more stressful. While this may sound like this analysis states that the general trend for Major League batters is to choke in clutch situations it could be that pitchers are actually performing better in clutch situations, or something else could be occurring entirely.

This analysis attempted to make a statement about individual clutch ability. However, chi-squared tests based on signs are plagued by an intricate bias that calls into question the results of the tests. Both tests would be very useful for determining if an individual clutch ability existed, but the bias issue would need to be resolved. The final round of tests found no evidence to reject the hypothesis that individuals do not have an inherent clutch ability. In conclusion, there is evidence that suggests that clutch batting averages are lower across all major league batters when compared to non-clutch batting averages; however, there is not enough evidence to show that certain individuals have better clutch abilities that others.

B. CLUTCH DEFINITIONS

There are many aspects of this analysis that could be examined in much greater detail. First, the definitions of clutch used in this analysis are based on easily measured aspects of the game at the time the batter is batting. As mentioned before, the batter could be stressed by other factors than those used in this analysis. Certainly there are some batters who fear getting sent to the minor leagues for bad hitting. This stressor would be extremely hard to measure and would take place at most any time the particular batter had an at-bat. Therefore, it would be hard to determine that particular batter's non-clutch batting average.

There are several ambient effects that would stress a batter as well. It is generally agreed that some playing fields favor pitchers and other playing fields favor batters. This is due to certain flexibility in the design of baseball parks with regards to fence heights. Also, some fields have consistent wind patterns that can either help or hurt batters. In addition to ill winds, batting at night might be considered more challenging too. All of these ambient effects could stress batters, and determining how much stress these effects place on batters would be difficult. The dataset from Retrosheet does provide the time of day and ballpark in which the particular at-bat occurred. The amount by which to weight these factors is debatable, but they might not be completely insignificant.

Another factor that might stress batters significantly more than the previously mentioned factors is championship games. At-bats during championship games are definitely more stressful than at-bats during regular season games. However, there could be at-bats during championship games that are more stressful than others. A possible extension of this analysis would be to utilize clutch levels. Clutch levels could be used to determine how stressful a particular situation is and then weight those situations accordingly. Finally, combining all the before mentioned factors, a nearly limitless number of possible clutch definitions could be examined.

C. ALPHAS

Alpha was used to correct for the overall effect of clutch situations in order to study individual changes in performance. The one sample two-sided t-tests done on the standardized batting average differences all had *p*-values of 0; this established that a general clutch effect existed because the distribution of clutch batting averages was not equal to the distribution of non clutch batting averages. As stated before, the general alpha we used could have been used to create the same situational alphas that are seen in the actual data. However, typically only five percent of our simulated alphas had greater standard deviations than our actual alphas. These situational alphas need to be studied further and perhaps situational alphas might need to be utilized. This will be challenging, because most batters will have clutch at-bats in many different situations, so the corrective factor might have to be determined for each individual batter.

Another modification to alpha can be made as well. Since league-wide batting averages are different from situation to situation, it could be the case that a particular batter bats disproportionately more in favorable situations than in unfavorable situations. Even worse, a particular batter could have a large proportion of his clutch at-bats in favorable situations, and then a large proportion of his non-clutch at-bats in unfavorable situations. This specific example would result in a batter whose clutch ability would be over estimated by this analysis. This analysis assumed that individual batters bat in equal situational proportions across their non-clutch and clutch at-bats. This assumption seems reasonable, but further study could prove the need to factor the proportion of a batter's clutch and non clutch at-bats that occur in each situation.

The previous modification could also examine the proportion of clutch and non-clutch at-bats each batter had in batter-friendly ballparks, pitcher-friendly ballparks, day games, night games, regular games, and championship games. If a particular batter has a high proportion of clutch at-bats in a pitcher-friendly park and a high proportion of non-clutch at-bats in a batter-friendly park, then his clutch ability would be

underestimated by this analysis. Additionally, there could be other field effects that need to be examined. Ultimately, an interesting further analysis would be to determine the correction for batters with large differences between their clutch situational proportions and their non-clutch situational proportions.

APPENDIX A: CLUTCH DEFINITIONS

```
> clutch.definition.1
function(data)
{
        #The inning is in the seventh or later, there are runners in
        #scoring position, and the score differential is less than
        #or equal to three
        out <- data$Inn >= 7
        out <- out & data$Runners != "Empty" & data$Runners != "1"
        out <- out & abs(data$VSc - data$HSc) <= 3
        return(out)
}
> clutch.definition.2
function(data)
{
        #The inning is in the fifth or later, there are runners in
        #scoring position, and the score differential is less than
        #or equal to four
        out <- data$Inn >= 5
        out <- out & data$Runners != "Empty" & data$Runners != "1"
        out <- out & abs(data$VSc - data$HSc) <= 4
        return(out)
}
> clutch.definition.3
function(data)
{
        #The inning is in the fifth or later, there are runners in
        #scoring position, the score differential is less than
        #or equal to three, and there are two outs
        out <- data$Inn >= 7
        out <- out & data$Runners != "Empty" & data$Runners != "1"
        out <- out & data$O == 2
        out <- out & abs(data$VSc - data$HSc) <= 3
        return(out)
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B: CLUTCH PLAYER TABLE FUNCTION

```
> player.clu.year
function(data, number)
{
        #This function applies the clutch definition to the data then creates a table
        #of unique batters who meet the number requirements of at-bats for the given
        #data. The clutch table contains the number of non-clutch hits, clutch hits,
        #non-clutch situations, and the number of clutch situations. Using these
        #numbers additional colums are added to include the difference, the variance,
        #the difference with alpha, the standardided difference, and the sign column.
        data$Batter <- as.factor(data$Batter)
        data$Clutch <- clutch.definition(data)
        tbl <- table(data$Batter, data$Clutch)
        tbl2 <- tbl[tbl[, "TRUE"] >= number,  ]
        events.for.my.guys <- data[is.element(data$Batter, tbl2[, 1])]
        events.for.my.guys <- data[is.element(data$Batter, dimnames(tbl2)[[1]]),  ]
        my.guys.pa <- table(events.for.my.guys$Batter)
        my.guys.pa.non <- table(events.for.my.guys$Batter[events.for.my.guys$Clutch ==
                FALSE])
        my.guys.pa.clu <- table(events.for.my.guys$Batter[events.for.my.guys$Clutch ==
                TRUE])
        my.guys.hit.non <- tapply((events.for.my.guys$Hit > 0)[events.for.my.guys$
                Clutch == FALSE], events.for.my.guys$Batter[events.for.my.guys$Clutch ==
                FALSE], sum)
        my.guys.hit.clu <- tapply((events.for.my.guys$Hit > 0)[events.for.my.guys$
                Clutch == TRUE], events.for.my.guys$Batter[events.for.my.guys$Clutch ==
                TRUE], sum)
        clu.table <- (data.frame(clutch.hits = my.guys.hit.clu, clutch.situations =
                my.guys.pa.clu, non.clutch.hits = my.guys.hit.non,
                non.clutch.situations = my.guys.pa.non))
        clu.table$Difference <- clu.table$non.clutch.hits/clu.table$
                non.clutch.situations - clu.table$clutch.hits/clu.table$
                clutch.situations
        clu.table <- clu.table[clu.table$Difference != "NA",  ]
        c1 <- clu.table$non.clutch.hits
        n1 <- clu.table$non.clutch.situations
        c2 <- clu.table$clutch.hits
        n2 <- clu.table$clutch.situations
        clu.table$Variance <- (((c1/n1) * (1 - c1/n1))/n1) + (((c2/n2) * (1 - c2/n2))/
                n2)
        clu.table <- clu.table[sign(clu.table$Variance) != 0,  ]
        c1 <- clu.table$non.clutch.hits
        n1 <- clu.table$non.clutch.situations
        c2 <- clu.table$clutch.hits
        n2 <- clu.table$clutch.situations
        clu.table$Standard.Difference <- (c1/n1 - c2/n2)/sqrt(clu.table$Variance)
        clu.table$AlphaDiff <- clu.table$Difference - alpha
        clu.table$sign <- sign(clu.table$AlphaDiff)
        return(clu.table[clu.table[, 2] >= number,  ])
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C: SIGN TABLE FUNCTION

```
> sign.table
function(year1, year2, year3, year4, year5, year6, year7, year8, cluAB)
{
        #Creates a table years and unique players who have meet the requirement of
        #clutch at bats for at least one of those years. The table is filled with
        #the values from the sign columns that correspond to the players in the
        #given year. A zero occurs when the batter shows up one year but not another.
        yearNames <- data.frame(year2000 = 0, year2001 = 0, year2002 = 0, year2003 = 0,
                year2004 = 0, year2005 = 0, year2006 = 0, year2007 = 0)
        sort(unique(c(row.names(year1[year1$clutch.situations >= cluAB,  ]), row.names(
                year2[year2$clutch.situations >= cluAB,  ]), row.names(year3[year3$
                clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ]))))
        size <- length(sort(unique(c(row.names(year1[year1$clutch.situations >= cluAB,
                ]), row.names(year2[year2$clutch.situations >= cluAB,  ]), row.names(
                year3[year3$clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ])))))
        cluab <- data.frame(matrix(0, size, 8))
        dimnames(cluab) <- list( + sort(unique(c(row.names(year1[year1$
                clutch.situations >= cluAB,  ]), row.names(year2[year2$
                clutch.situations >= cluAB,  ]), row.names(year3[year3$
                clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ])))),  + names(yearNames))

        cluab[row.names(year1[year1$clutch.situations >= cluAB,  ]),  ]$year2000 <-
                year1[year1$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year2[year2$clutch.situations >= cluAB,  ]),  ]$year2001 <-
                year2[year2$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year3[year3$clutch.situations >= cluAB,  ]),  ]$year2002 <-
                year3[year3$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year4[year4$clutch.situations >= cluAB,  ]),  ]$year2003 <-
                year4[year4$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year5[year5$clutch.situations >= cluAB,  ]),  ]$year2004 <-
                year5[year5$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year6[year6$clutch.situations >= cluAB,  ]),  ]$year2005 <-
                year6[year6$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year7[year7$clutch.situations >= cluAB,  ]),  ]$year2006 <-
                year7[year7$clutch.situations >= cluAB,  ]$sign
        cluab[row.names(year8[year8$clutch.situations >= cluAB,  ]),  ]$year2007 <-
                year8[year8$clutch.situations >= cluAB,  ]$sign
        big.ugly.signs <- cluab
        return(big.ugly.signs)
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D: ALL YEARS CLUTCH TABLE

```
> sum.clutch.year
function(data1, data2, data3, data4, data5, data6, data7, data8, number)
{
        #Applies the clutch definition to all the data sets. Smaller tables are
        #made for each of the years. All the unique players are then placed in
        # a larger table and the the number of non-clutch hits, non-clutch
        #situations, clutch hits, and clutch situations are added across
        #all years. Additional colums are added in the same way they were added
        #in the player.clu.year function
        data1$Clutch <- clutch.definition(data1)
        data2$Clutch <- clutch.definition(data2)
        data3$Clutch <- clutch.definition(data3)
        data4$Clutch <- clutch.definition(data4)
        data5$Clutch <- clutch.definition(data5)
        data6$Clutch <- clutch.definition(data6)
        data7$Clutch <- clutch.definition(data7)
        data8$Clutch <- clutch.definition(data8)
        cluab1 <- player.clu.year(data1, number)
        cluab2 <- player.clu.year(data2, number)
        cluab3 <- player.clu.year(data3, number)
        cluab4 <- player.clu.year(data4, number)
        cluab5 <- player.clu.year(data5, number)
        cluab6 <- player.clu.year(data6, number)
        cluab7 <- player.clu.year(data7, number)
        cluab8 <- player.clu.year(data8, number)
        sort(unique(c(row.names(cluab1), row.names(cluab2), row.names(cluab3),
                row.names(cluab4), row.names(cluab5), row.names(cluab6), row.names(
                cluab7), row.names(cluab8))))
        size <- length(sort(unique(c(row.names(cluab1), row.names(cluab2), row.names(
                cluab3), row.names(cluab4), row.names(cluab5), row.names(cluab6),
                row.names(cluab7), row.names(cluab8)))))
        cluab <- data.frame(matrix(0, size, 9))
        dimnames(cluab) <- list( + sort(unique(c(row.names(cluab1), row.names(cluab2),
                row.names(cluab3), row.names(cluab4), row.names(cluab5), row.names(
                cluab6), row.names(cluab7), row.names(cluab8)))),  + names(cluab1))
        cluab[row.names(cluab1),  ] <- cluab1
        cluab[row.names(cluab2),  ] <- cluab[row.names(cluab2),  ] + cluab2
        cluab[row.names(cluab3),  ] <- cluab[row.names(cluab3),  ] + cluab3
        cluab[row.names(cluab4),  ] <- cluab[row.names(cluab4),  ] + cluab4
        cluab[row.names(cluab5),  ] <- cluab[row.names(cluab5),  ] + cluab5
        cluab[row.names(cluab6),  ] <- cluab[row.names(cluab6),  ] + cluab6
        cluab[row.names(cluab7),  ] <- cluab[row.names(cluab7),  ] + cluab7
        cluab[row.names(cluab8),  ] <- cluab[row.names(cluab8),  ] + cluab8
```

```
        cluab$Difference <- cluab$non.clutch.hits/cluab$non.clutch.situations - cluab$
                clutch.hits/cluab$clutch.situations
        cluab <- cluab[cluab$Difference != "NA",  ]
        c1 <- cluab$non.clutch.hits
        n1 <- cluab$non.clutch.situations
        c2 <- cluab$clutch.hits
        n2 <- cluab$clutch.situations
        cluab$Variance <- (((c1/n1) * (1 - c1/n1))/n1) + (((c2/n2) * (1 - c2/n2))/
                n2)
        cluab <- cluab[sign(cluab$Variance) != 0,  ]
        c1 <- cluab$non.clutch.hits
        n1 <- cluab$non.clutch.situations
        c2 <- cluab$clutch.hits
        n2 <- cluab$clutch.situations
        cluab$Standard.Difference <- (c1/n1 - c2/n2)/sqrt(cluab$Variance)
        cluab$AlphaDiff <- cluab$Difference - alpha
        cluab$sign <- sign(cluab$AlphaDiff)
        return(cluab[cluab[, 2] >= number,  ])
        return(cluab)
}
```

APPENDIX E: CHI-SQUARED ON THE BINOMIAL DISTRIBUTION FUNCTION

```
> binomial.chisq
function(number)
{
        #The number passed in is the clutch at-bat requirement. The
        #sign table is made, but a smaller table is made of batters
        #who appreared in at least four years.The zero's are removed
        #and that compressed table is passed to the chisq.four.unif.years
        big.ugly.signs <- sign.table(player.clu.year(ab.00, number), player.clu.year(
                ab.01, number), player.clu.year(ab.02, number), player.clu.year(
                ab.03, number), player.clu.year(ab.04, number), player.clu.year(
                ab.05, number), player.clu.year(ab.06, number), player.clu.year(
                ab.07, number), number)
        abstabs <- abs(big.ugly.signs)
        abstabs$sums <- abstabs$year2000 + abstabs$year2001 + abstabs$year2002 +
                abstabs$year2003 + abstabs$year2004 + abstabs$year2005 + abstabs$
                year2006 + abstabs$year2007
        big.ugly.signs$sums <- abstabs$sums
        small.ugly.signs <- big.ugly.signs[big.ugly.signs$sums >= 4,  ]
        copy <- small.ugly.signs
        length(row.names(small.ugly.signs))
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                 else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                 else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                 else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                 else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                 else (copy[i, j - 1] <- 0)
```

```
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.signs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                        if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                                else (copy[i, j - 1] <- 0)
        chi.table <- copy[, 5:8]
        chi.table.4 <- copy[, 5:8]
        for(i in 1:length(row.names(chi.table)))
                for(j in 1:4)
                        if(chi.table.4[i, j] == -1) chi.table.4[i, j] <- 0
        chi.table.4$sums <- chi.table.4$year2004 + chi.table.4$year2005 +
                chi.table.4$year2006 + chi.table.4$year2007
        binomtbl <- table(chi.table.4$sums)
        binomdata <- rep(as.numeric(names(binomtbl)), binomtbl)
        (chisq.gof(binomdata,  , seq(-0.5, 4.5, by = 1), dist = "binomial", size = 4,
                prob = 0.5, n.param.est = 0))
        return(chi.table.4)
}
```

APPENDIX F: CHI-SQUARED FOR THE UNIFORM DISTRIBUTION FUNCTION

```
> chisq.unif.four.years
function(FourYearTable)
{
        #The FourYearTable is the compressed table made in the
        #binomial.chisq function ater all the zero are removed.
        copy.table <- FourYearTable
        copy.signtbl <- table(apply(copy.table, 1, function(x)
        paste(signs(unlist(x)), collapse = "")))
        test <- chisq.for.discrete.unif(copy.signtbl)
        return(copy.signtbl, test)
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX G: DIFFERENCE TABLE FUNCTION

```
> diff.table
function(year1, year2, year3, year4, year5, year6, year7, year8, cluAB)
{
        #Creates a table years and unique players who have meet the requirement of
        #clutch at bats for at least one of those years. The table is filled with
        #the standardized alpha difference values that correspond to the players in the

        #given year. A zero occurs when the batter shows up one year but not another.
        yearNames <- data.frame(year2000 = 0, year2001 = 0, year2002 = 0, year2003 = 0,
                year2004 = 0, year2005 = 0, year2006 = 0, year2007 = 0)
        sort(unique(c(row.names(year1[year1$clutch.situations >= cluAB,  ]), row.names(
                year2[year2$clutch.situations >= cluAB,  ]), row.names(year3[year3$
                clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ]))))
        size <- length(sort(unique(c(row.names(year1[year1$clutch.situations >= cluAB,
                ]), row.names(year2[year2$clutch.situations >= cluAB,  ]), row.names(
                year3[year3$clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ])))))
        cluab <- data.frame(matrix(0, size, 8))
        dimnames(cluab) <- list( + sort(unique(c(row.names(year1[year1$
                clutch.situations >= cluAB,  ]), row.names(year2[year2$
                clutch.situations >= cluAB,  ]), row.names(year3[year3$
                clutch.situations >= cluAB,  ]), row.names(year4[year4$
                clutch.situations >= cluAB,  ]), row.names(year5[year5$
                clutch.situations >= cluAB,  ]), row.names(year6[year6$
                clutch.situations >= cluAB,  ]), row.names(year7[year7$
                clutch.situations >= cluAB,  ]), row.names(year8[year8$
                clutch.situations >= cluAB,  ]))))),  + names(yearNames))
```

```
        cluab[row.names(year1[year1$clutch.situations >= cluAB,  ]),  ]$year2000 <-
                year1[year1$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year2[year2$clutch.situations >= cluAB,  ]),  ]$year2001 <-
                year2[year2$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year3[year3$clutch.situations >= cluAB,  ]),  ]$year2002 <-
                year3[year3$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year4[year4$clutch.situations >= cluAB,  ]),  ]$year2003 <-
                year4[year4$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year5[year5$clutch.situations >= cluAB,  ]),  ]$year2004 <-
                year5[year5$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year6[year6$clutch.situations >= cluAB,  ]),  ]$year2005 <-
                year6[year6$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year7[year7$clutch.situations >= cluAB,  ]),  ]$year2006 <-
                year7[year7$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        cluab[row.names(year8[year8$clutch.situations >= cluAB,  ]),  ]$year2007 <-
                year8[year8$clutch.situations >= cluAB,  ]$StandAlpha.Difference
        big.ugly.diffs <- cluab
        return(big.ugly.diffs)
}
```

APPENDIX H: CONSECUTIVE YEARS DIFFERENCE TABLE FUCNTION

```
> diff.tabler
function(number, years)
{
        #This function produces a large standardized difference table
        #then grabs only the gy=uys who appear in at least four years.
        #The zeros are then removed and that new table is returned.
        big.ugly.diffs <- diff.table(player.clu.year(ab.00, number), player.clu.year(
                ab.01, number), player.clu.year(ab.02, number), player.clu.year(ab.03,
                number), player.clu.year(ab.04, number), player.clu.year(ab.05, number),
                player.clu.year(ab.06, number), player.clu.year(ab.07, number), 1)
        abstabs <- abs(sign(big.ugly.diffs))
        abstabs$sums <- abstabs[, 1] + abstabs[, 2] + abstabs[, 3] + abstabs[, 4] +
                abstabs[, 5] + abstabs[, 6] + abstabs[, 7] + abstabs[, 8]
        big.ugly.diffs$sums <- abstabs$sums
        small.ugly.diffs <- big.ugly.diffs[big.ugly.diffs$sums >= years,  ]
        copy <- small.ugly.diffs
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
        for(i in 1:length(row.names(small.ugly.diffs)))
                for(j in 8:2)
                        if(copy[i, j] == 0) for(g in 0:1)
                                if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
                                        else (copy[i, j - 1] <- 0)
```

```
	for(i in 1:length(row.names(small.ugly.diffs)))
		for(j in 8:2)
			if(copy[i, j] == 0) for(g in 0:1)
					if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
						else (copy[i, j - 1] <- 0)
	for(i in 1:length(row.names(small.ugly.diffs)))
		for(j in 8:2)
			if(copy[i, j] == 0) for(g in 0:1)
					if(g == 0) copy[i, j] <- copy[i, ((j) - 1)]
						else (copy[i, j - 1] <- 0)
	return(copy)
}
```

APPENDIX I: RANKING FUNCTION

```
> the.ranker
function(hope)
{
        #This funtion ranks each batter in each year of the four
        #year table. After the batters are ranked they are placed
        #in quartiles for each year.
        the.ranks <- hope
        for(i in 1:4)
                the.ranks[, i] <- (rank(hope[, i]))
        maxRank <- max(the.ranks$year2004)
        for(i in 1:4)
                (the.ranks[the.ranks[, i] <= maxRank/4,  ][, i] <- 1)
        for(i in 1:4)
                (the.ranks[the.ranks[, i] > maxRank/4 & the.ranks[, i] <= (
                        2 * maxRank)/4,  ][, i] <- 2)
        for(i in 1:4)
                (the.ranks[the.ranks[, i] > (2 * maxRank)/4 & the.ranks[, i] <=
                        (3 * maxRank)/4,  ][, i] <- 3)
        for(i in 1:4)
                (the.ranks[the.ranks[, i] > (3 * maxRank)/4,  ][, i] <- 4)
        return(the.ranks)
}
```

THIS PAGE INTENTIONALLY LEFT BLANK

BIBLIOGRAPHY

Conover, W.J. *Practical Nonparametric Statistics*. New York: John Wiley & Sons Inc, 1999.

Cramer, Richard D. "Do Clutch Hitters Exist?" Baseball Research Journal (1977), http://www.geocities.com/cyrilmorong@sbcglobal.net/CramerClutch2.htm (accessed April 12, 2008).

Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. California: Thompson Corporation, 2008.

Fuld, Elan. "Clutch and choke Hitters in Major League Baseball: Romantic Myth or Empirical Fact." 1st Draft (May 19, 2005).

Grabiner, David. Do Clutch Hitters Exist? Paper presented at the SABRBoston Presents Sabermetrics conference, May 20, 2006.

Hakes, Jahn H., and Raymond D. Sauer. "Are players Paid for 'Clutch' Performance?" John E. Walker Dept. of Economics, Clemson University. Preliminary Draft (June 20, 2003), http://people.albion.edu/jhakes/pdfs/clutch.pdf (accessed March 18, 2008).

Retrosheet. http://retrosheet.org/ (accessed December 10, 2007).

Ruane, Tom. "In Search of Clutch Hitters." Baseball Research Journal (January 1, 2005), http://retrosheet.org/Research/RuaneT/clutch_art.htm (accessed March 17, 2008).

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Fort Belvoir, VA

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, CA

3. Daniel Ammons-Moreno
   Hendersonville, NC

4. Kathleen Pauls
   LaFayette, NY

5. Jerome Richman
   Crofton, MD

6. Judson C. Ammons
   Hendersonville, NC

7. Samuel E. Buttrey
   Naval Postgraduate School
   Monterey, CA

8. David W. Meyer
   Naval Postgraduate School
   Monterey, CA